

IMPROVING SIGNAL RESOLUTION AND REDUCING EXPERIMENT TIME IN  
ELECTRON SPIN RESONANCE SPECTROSCOPY VIA DATA PROCESSING  
METHODS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Madhur Srivastava

August 2018

© 2018 Madhur Srivastava

IMPROVING SIGNAL RESOLUTION AND REDUCING EXPERIMENT TIME IN  
ELECTRON SPIN RESONANCE SPECTROSCOPY VIA DATA PROCESSING  
METHODS

Madhur Srivastava, Ph. D.

Cornell University 2018

Electron Spin Resonance (ESR) Spectroscopy is a powerful method to study structure and dynamics of biomolecules. It detects magnetic resonance signals from the spin labels (or paramagnetic tags) attached at specific locations of a biological sample such as proteins, DNA, RNA, and lipids, among others. Despite many improvements in ESR instrument sensitivity, experimental signals lack sufficient strength that is needed to study complex biomolecules. We introduce signal processing methods as an alternate approach to remove noise present in the experimental ESR signals. We developed new methods based on wavelet transforms that are effectively able to retrieve signals, including weak signals. The methods are developed for one-dimensional continuous wave electron spin resonance (cw-ESR) signals and pulsed dipolar signals. The denoising methods achieve more than two orders-of-magnitude improvement in Signal-to-Noise Ratio (SNR) as well as can be used to reduce signal acquisition time by similar orders.

In cw-ESR, the experimental signals obtained from the spectrometer are directly analyzed. Hence, there is no need for further data processing. In pulsed dipolar spectroscopy, the acquired signal needs to be processed to obtain desired information, i.e. distance distribution between a pair of spin labels. In addition to wavelet

denoising, we developed a singular value decomposition (SVD) method to obtain well-resolved accurate distance distributions from the dipolar signals. The SVD method overcomes the limitations of standard Tikhonov regularization method that yields are compromise between a good resolution and a stable solution, as well as the limitation of model fitting methods that require *a priori* information. Further, we conducted the uncertainty analysis of the SVD solution.

This research is first step towards utilizing the power of data processing methods to improve signal analysis in ESR spectroscopy.

## BIOGRAPHICAL SKETCH

Madhur Srivastava was born (December 29, 1988) and raised in Lucknow, India. Initially interested in being a professional cricketer, he decided to pursue an undergraduate degree in engineering after taking into account his incompetence in the sport. He received the B.Tech degree in Electronics and Communication Engineering from Jaypee University of Engineering and Technology (JUET), Guna, India, in 2011. Following which, he came to Cornell University, Ithaca, NY, USA to pursue the M.Eng. degree in Electrical and Computer Engineering and graduated in 2012. Subsequently, he worked as a research assistant for one and a half years. He later joined the Ph.D. program in the Meinig School of Biomedical Engineering at Cornell University in January 2014 and became graduate research assistant at the National Biomedical Center for Advanced Electron Spin Resonance Technology (ACERT) under the supervision of Professor Jack H. Freed. During the program, he took minors in Biological Engineering and Biophysics and received the M.S. degree in Biomedical Engineering after passing his candidacy exam in 2016. Apart from formal education, he did a summer internship at the Indian Institute of Science Education and Research-Kolkata, Nadia, India in 2010, for which he was also awarded a fellowship. In 2015, he did a summer immersion program at Weill Cornell Medicine, New York, NY, USA.

He has been a student member of the International Electron Paramagnetic Resonance Society, American Chemical Society, Biophysical Society, International Society of

Magnetic Resonance, New York Academy of Sciences, and Institution of Electronics and Telecommunications Engineers (IETE).

He is a recipient of many awards and fellowships that include the Vice-Chancellor's Gold Medal at JUET (2011), the Ph.D. Technology Commercialization Fellowship (2016) and the Phool Prakash and Rukmini Sahai Graduate Fellowship (2017), both of them at Cornell University, the Electron Spin Resonance Technology Transfer Experience Award (2017) by SharedEPR, the Alumni Association Academic Enrichment Grant (2017) and Graduate School Research Grant (2017) by Cornell University. He received conference travel awards from the Graduate School of Cornell University, the 20<sup>th</sup> International Society of Magnetic Resonance Meeting (2017), the International Conference on EPR Spectroscopy and Imaging of Biological Systems (2017), and the 58<sup>th</sup> Rocky Mountain Conference on Magnetic Resonance (2016). He was featured in the Graduate School Student Spotlight (2016), the Broadening Experiences in Scientific Training Program Student Spotlight (2017), the Cornell Chronicle, The Cornell Daily Sun, and the Health Imaging.

Dedicated to my family members and mentors, and to my little nephew Sarvagya

## ACKNOWLEDGMENTS

Human existence depends on cooperation and mutual support. Born in the species of *Homo sapiens*, I benefitted from such genes and environment that has led to this dissertation. My Ph.D. research would not have been possible without mentorship, guidance, encouragement and assistance from many individuals. I have been lucky to have them in my life.

**Family Members:** First, I would like to thank my parents, Dr. Rajendra Srivastava and Jyoti Srivastava, who raised me in a loving environment and provided constant support to pursue my dreams. Despite having contracted all the possible diseases a child can have, I always felt healthy, both mentally and physically, due to the care and attention I received from my parents. My father's life has been an inspiration to me that has led me to pursue my life, research and career goals. In spite of being a senior government official, he passionately volunteered to teach classes in civil engineering and designed school buildings without any fees. I saw many of his students staying at our home to study and work on real projects. This provided me initial academic environment with strong social motivation. Always passionate about his work, my father pursued his Ph.D. in his early 50s. This determination and perseverance inspired me to follow my goals, irrespective of time, age and circumstances. My mother, on the other hand, always provided me the emotional support, especially when I consistently failed in courses throughout my schools years. My father never let success to my head and my mother never let failure to my heart. I learned from my father that a stone that

is afraid of hammer can never become a sculpture.

I would also like to thank my brother, Manas Srivastava, from whom I learned how to be humble and considerate to everyone. Although he never felt shy of breaking my bones (or head), he ensured that I was protected when needed. He also ensured that I was not the son who received the lowest marks, boosting my morale.

**Ph.D. Advisor:** My Ph.D. advisor and committee chair, Prof. Jack Freed, has been a constant source of motivation for me. He brings out the best in me. I have learned from him how to convert abstract ideas into useful and meaningful research. Conversations with him are so enriching and informative that I come up with new ideas regularly, and love to do research. Coming from India, I was looking for an advisor who not only would provide me with guidance in academics and research, but also would be a life mentor. I also wanted an advisor who is always accessible and with whom I can share new ideas any time. My advisor, Professor Jack Freed, likes to be closely involved with my research. He facilitated discussion and allows me to take the lead. He likes intellectual stimulation and enjoys conversations. One particularly important thing which I found about him is that he doesn't make you wait for him. If I provide a research paper draft or ask a question to him, he instantly acts on it. Knowing that I am an international student, he was cognizant of my more limited English language skills and helps me improve it by helping me to correct grammatical and other errors. I sought active collaboration with research groups, many of them outside my lab's research area. Professor Freed has been an active facilitator in these

endeavors. In general, he has a very positive outlook on things and never gives up easily. He also focuses on details, which I realize is very important in making things work as well as standing the test of time. An extra second spent on a task can save a thousand minutes later. He writes very detailed and comprehensive papers, a skill I am learned from him to my benefit. Another thing that I learned from him is how to use time effectively: Is it worth spending the extra time considering the outcome we hope to accomplish?

While conducting research, there are many times when things don't go according to plan. My advisor always has a story and/or experience to share which helps me realize that the situation is just part and parcel of being a researcher, and things will eventually work out. Also, I realized that I am not the only one who encounters such situations! He has taught me how to identify and capitalize on my strengths and how to manage/work on my weaknesses. For instance, lacking good research writing skills, it takes me more research paper drafts than for others. He has been instrumental in making me start early. Another aspect I learned from him is writing research grants. Grant writing is more an art than it is science! One must be able to effectively communicate to the reviewers within the guidelines and page limits.

**Ph.D. Committee Members:** I would like to express my gratitude towards Professor Gerald Feigenson and Professor C. Lindsay Anderson for providing constant support and feedback on my research, and meeting me regularly during the Ph.D. program. Through them, I have learned how to ask important questions about my research, both

from within a specific field and a general point of view. They provided me guidance on career development and actively advised me on how to approach a career in academia.

**Previous Mentors:** I have been lucky to have great mentors in my life from whom I sought both inspiration and guidance. I learned from my undergraduate physics professor, Dr. Rajneesh Atre, to enjoy research and to work selflessly. My undergraduate supervisor, Professor Satish Singh, trained and prepared me for graduate school. Professor Prasanta Panigrahi, under whom I did a summer internship, cultivated in me scientific curiosity and openness toward new ideas. My M.Eng. supervisor, Professor Sheila Hemami, taught me critical thinking, and how to efficiently and sensibly conduct research.

**ACERT/Lab Members:** I was grateful to have friendly, collaborative, and fun environment that I share with my lab members. The environment enhanced my research capabilities. Amount of experimental data sets I received to test my research methods was unprecedented.

**Friends and Flatmates:** Exchange of ideas and lively academic conversations outside lab with my friends Siddarth Chandsekar, Amandeep Gupta, Aritro Sinharoy, Aadhar Jain, Munir Ahmad Nayak, and Poornima Padmanabhan have greatly enhanced my scholarship. Because of Cornell, I am fortunate to have flatmates like Tayler Eaton, Abby Yip, Enoch Kwon, and Mate Rigo, who come from different

countries, cultures, and academic backgrounds. They influenced me to think not just in terms of science and engineering, but also in terms of history, society, and community as a whole.

**Dickson and McLLU Staff:** During my graduate studies, I got a unique opportunity to be the Assistant Residence Hall Director of a freshman dorm, Dickson Hall, where I worked with my boss Jamie Hom and 16 Resident Advisors. The experience of running a dorm house with them inculcated community living in me. Working with them was one of the best experiences of my stay at Cornell. It also made my research more productive.

**Graduate Field Assistants (GFA):** I want to convey my gratitude towards Belinda Floyd and Pat Hine, the GFAs of Biomedical Engineering and Chemistry, respectively, who made my Ph.D. hassle-free and ensure that I got all support I need from the department.

## TABLE OF CONTENTS

Chapter 1: Signal Analysis and Denoising: Limitations of Fourier Transform and Short Time Fourier Transform, and Introduction to Wavelet Transform.....	1
1.1 Fourier Transform.....	2
1.1.1 Limitations of Frequency Analysis using FT.....	6
1.1.2 Effects of Noise Presence on Signal Information in Frequency Domain.....	6
1.1.3 Frequency-based Noise Filtering and its Limitations.....	9
1.2 Short Time Fourier Transform.....	13
1.2.1 Limitations of STFT.....	13
1.2.2 Limitations of STFT-based Noise Filtering.....	17
1.3 Wavelet Transform.....	17
1.3.1 Heisenberg Uncertainty Boxes for Signal Domain-Frequency Resolution.....	19
1.3.2 Wavelet Selection.....	20
1.4 Discrete Wavelet Transform.....	23
1.4.1 Introduction to Approximation Component.....	27
1.4.2 Scaling Function and its relation to Wavelet Function.....	29
1.4.3 Inverse Discrete Wavelet Transform.....	32
1.4.4 Relationship between $j^{th}$ Approximation Component, and $j + 1^{th}$ Approximation and Detail Component.....	32
1.5 Fast Wavelet Transform.....	34
1.5.1 Relationship between Low and High Pass Filters.....	40
1.5.2 Orthogonal and Biorthogonal Wavelets.....	42
1.6 Wavelet Denoising – Noise Removal using Wavelet Transforms.....	44

1.6.1 Wavelet Denoising Procedure.....	47
References.....	49
Chapter 2: A New Wavelet denoising method for selecting decomposition levels and noise thresholds.....	50
Abstract.....	50
Graphical Abstract.....	51
2.1 Introduction.....	52
2.2 Wavelet Shrinkage Denoising Method.....	53
2.2.1 Algorithm: Wavelet Shrinkage Denoising.....	55
2.2.2 Limitations of Current Shrinkage Methods.....	57
2.3 Noise in Wavelet Domain.....	58
2.4 Wavelet Bases.....	60
2.5 New Denoising Method.....	61
2.5.1 Decomposition Level Selection.....	63
2.5.1.1 Subjective Method.....	63
2.5.1.2 Objective Method.....	66
2.5.2 Threshold Selection.....	67
2.5.3 Thresholding Function.....	72
2.5.4 Denoising the Approximation Component.....	73
2.6 ESR Experiments and Results.....	74
2.6.1 ESR Methodology and Sample Preparation.....	74
2.6.2 Experiment.....	75
2.6.3 Results.....	78
2.6.4 Computational Complexity.....	89
2.7 Conclusions.....	91

References.....	92
Chapter 3: Noise elimination and Reduction via denoising: A wavelet approach for 1-D spectroscopy.....	97
Abstract.....	97
Graphic Abstract.....	98
3.1 Introduction.....	99
3.2 Methods and Experiments.....	102
3.2.1 Wavelet Transforms and Discrete Wavelet Transforms.....	102
3.2.2 Wavelet Denoising.....	105
3.2.2.1 Undecimated Discrete Wavelet Transform.....	105
3.2.2.2 Wavelet Selection.....	109
3.2.2.3 Decomposition Level Selection.....	109
3.2.2.4 Noise Thresholds Calculation.....	110
3.2.2.5 Noise Thresholding Function.....	111
3.2.2.6 Denoising Approximation Component.....	112
3.2.3 Difference between NERD and Standard Denoising.....	113
3.2.4 Spectroscopy Method.....	117
3.2.5 Model Data Generation.....	117
3.2.6 Experimental Data Generation.....	117
3.2.6.1 ESR Experiment and Sample Preparation.....	117
3.2.6.2 Signal Averaging.....	118
3.2.7 Objective Measure.....	119
3.3 Results.....	119
3.3.1 Model Data.....	119
3.3.2 Experimental Data.....	123
3.3.2.1 Example 1: Non-overlapping signals.....	123

3.3.2.2 Example 2: Overlapping signals.....	126
3.4 Discussion.....	129
3.5 Conclusion.....	129
References.....	131
Chapter 4: A New Wavelet Denoising Method for Experimental Time-domain Signals: Pulsed Dipolar Electron Spin Resonance.....	135
Abstract.....	135
Graphical Abstract.....	136
4.1 Introduction.....	137
4.1.1 Background.....	137
4.1.2 Conversion of Pulse Dipolar Signals to Distance Distributions .....	139
4.1.3 Wavelet Transform and Discrete Wavelet Transform.....	144
4.2 Method.....	147
4.2.1 Wavelet Denoising.....	147
4.2.2 Difference between the New Method and Standard Methods...	153
4.2.3 Spurious Peaks.....	153
4.2.4 Objective Measures.....	156
4.2.5 ESR Method and Sample Preparation.....	157
4.3 Results.....	158
4.3.1 Model Data.....	158
4.3.1.1 Unimodal Distance Distribution.....	158
4.3.1.2 Bimodal Distance Distribution.....	164
4.3.1.3 Short Evolution Time vs. Long Evolution Time.....	168
4.3.2 Experimental Data.....	171
4.3.2.1 Unimodal Case.....	171

4.3.2.2 Bimodal Case.....	175
4.4 Discussion.....	178
4.5 Conclusion.....	185
4.S Supplementary Information.....	186
References.....	207
Chapter 5: Singular Value Decomposition Method to Determine Distance Distributions in Pulsed Dipolar Electron Spin Resonance.....	214
Abstract.....	214
Graphical Abstract.....	215
5.1 Introduction.....	216
5.2 Pulsed Dipolar Spectroscopy.....	217
5.3 Current Methods for Distance Distribution Reconstruction.....	218
5.4 SVD-based Method.....	219
5.5 Results and Discussion.....	231
5.5.1 Experimental and Model Data.....	231
5.5.2 WavPDS.....	232
5.5.3 Experimental Case Study.....	232
5.6 Conclusion.....	234
5.S Supplementary Information.....	235
References.....	261
Chapter 6: Singular Value Decomposition Method to Determine Distance Distributions in Pulsed Dipolar Electron Spin Resonance: II. Estimating Uncertainty.....	266
Abstract.....	266
Graphical Abstract.....	267
6.1 Introduction.....	268

6.2 Method.....	269
6.2.1 Pulsed Dipolar Spectroscopy.....	269
6.2.2 Distance Distribution Reconstruction in PDS using SVD.....	269
6.2.3 Uncertainty Determination.....	271
6.2.3.1 Determination of Singular Value Cut-off.....	271
6.2.3.2 Identification of Minimum and Maximum SVCs.....	272
6.2.3.3 Calculation of Minimum and Maximum Value of the Distance in the Min-Max SVC Range.....	273
6.3 Results and Discussion.....	276
6.3.1 Experimental Data.....	276
6.3.2 Distance Reconstruction and Uncertainty Analysis.....	277
6.4 Conclusion.....	282
References.....	284
Chapter 7: Conclusion.....	287
7.1 Wavelet Denoising.....	287
7.2 Distance Reconstruction in Pulsed Dipolar Spectroscopy.....	288
Appendix 1: Wavelet Filters: Filter Values and Matlab Script.....	290

# CHAPTER 1

## SIGNAL ANALYSIS AND DENOISING: LIMITATIONS OF FOURIER TRANSFORM AND SHORT TIME FOURIER TRANSFORM, AND INTRODUCTION TO WAVELET TRANSFORM

*Wavelet Transform is in 21<sup>st</sup> Century what Fourier Transform was in 20<sup>th</sup> Century – Jack H. Freed*

A transformation of a signal is usually required to analyze or extract a particular characteristic/feature present in an experimental signal. The frequency information of a signal is the most commonly used feature for a variety of applications. A signal, in general, is a collection of period functions having definite frequency or frequencies. Thus, the frequency information can easily reveal the characteristics of an individual period function, which is hidden in the signal domain.

*An arbitrary function, continuous or with discontinuities, defined in a finite interval by an arbitrarily capricious graph can always be expressed as a sum of sinusoids – J.B.J. Fourier*

## 1.1 Fourier Transform

The Fourier transform (FT) is most widely used to analyze a time-domain signal. It provides the frequency information of a signal that represents the frequencies and their respective energy density present in the signal. The FT of a signal is defined as<sup>1</sup>,

$$F(\omega) = \int_{-\infty}^{+\infty} f(t)e^{-i\omega t} dt \tag{1.1}$$

where  $f(t)$  and  $F(\omega)$  are the input signal and the FT, respectively. The input signal is usually in the time domain but can be in an other domain, such as magnetic field sweep as in continuous wave – Electron Spin Resonance (cw-ESR) spectroscopy<sup>2</sup>.

The FT is the invertible transform and the input signal can be reconstructed by<sup>1</sup>,

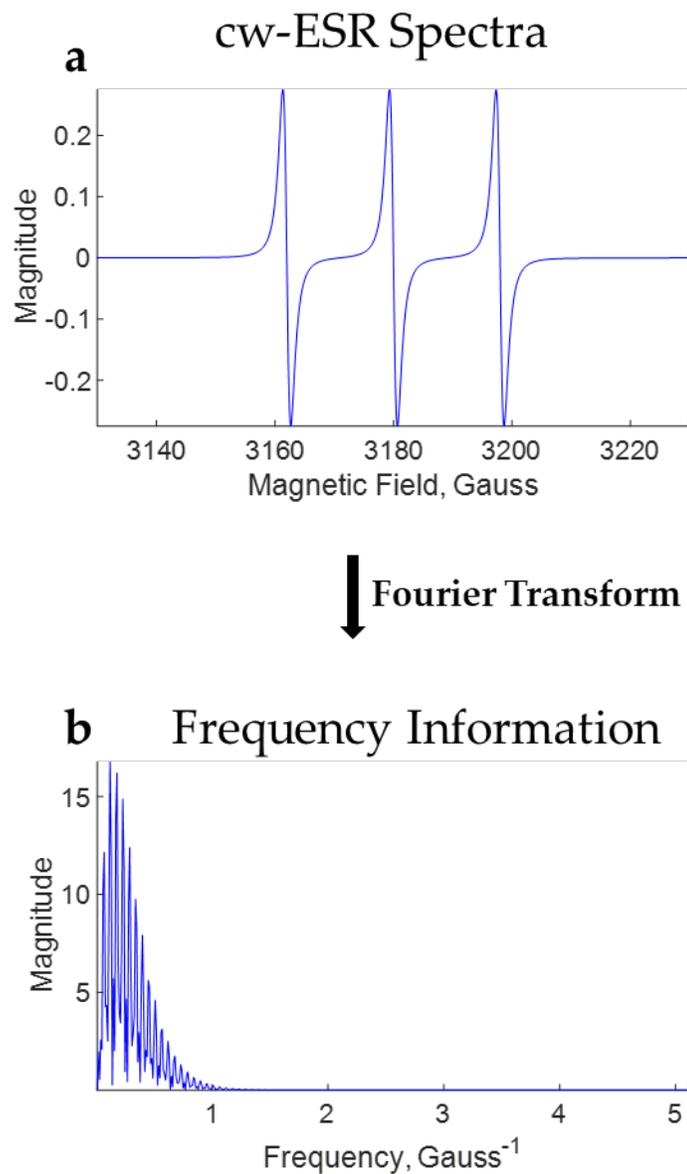
$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega)e^{i\omega t} d\omega \tag{1.2}$$

Figs. 1.1 and 1.2 show the frequency information of cw-ESR and pulsed dipolar (model/synthetic) signal using the FT, respectively. The Fast Fourier Transform (FFT) method was applied on discrete signals to obtain the FT. The cw-ESR spectra is acquired in magnetic field domain and is represented in Gauss. On the other hand, the

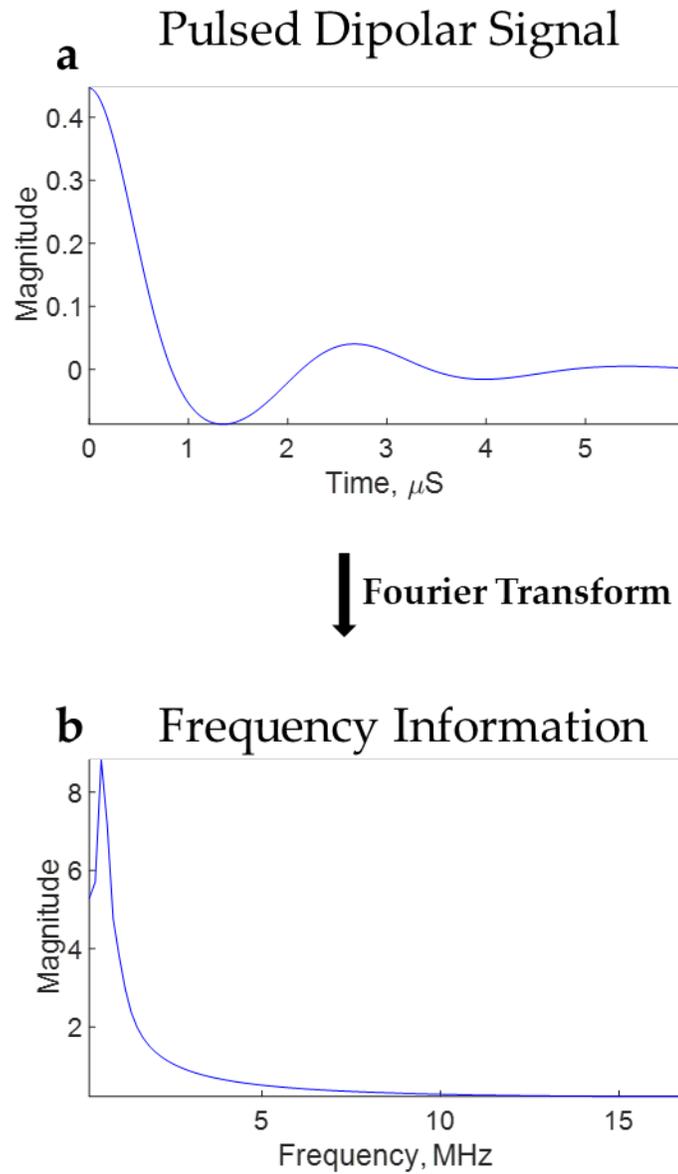
pulsed dipolar signal is a time domain signal (which is standard for most experimental signals).

The model cw-ESR signal was generated using Lorentzian functions and the model pulsed dipolar signal was generated using the formula  $S = KP$ , where  $K$  is kernel matrix obtained from Fredholm equation and  $P$  is the unimodal Gaussian distance distribution centered at  $5 \text{ nm}$  with standard deviation of  $0.3$ .

The FT reveals that the cw-ESR dipolar signal, in Fig. 1.1, has predominantly low frequencies in which energy density reduces with the increase in frequency range. After  $1.5 \text{ G}^{-1}$ , there is no higher frequency component present in the cw-ESR signal. Similarly, the pulsed dipolar signal (cf. Fig. 1.2) is dominated by lower frequencies, though the frequency range is very high (in  $\text{MHz}$ ). It is impossible to extract such information in the signal domain.



**Figure 1.1:** Frequency information of the model/synthetic cw-ESR signal using the FT. **a)** Input cw-ESR spectra; and **b)** Fourier-transformed spectra. The cw-ESR spectra is measured in Gauss (magnetic field) and hence the frequency information is in Gauss<sup>-1</sup> (G<sup>-1</sup>).



**Figure 1.2:** Frequency information of the model/synthetic pulsed dipolar signal using the FT. **a)** Input pulsed dipolar signal; and **b)** Fourier-transformed spectrum.

### **1.1.1 Limitations of Frequency Analysis using FT**

A major limitation of the FT is that it can only yield the frequency information of the complete signal. It lacks the capability to provide frequency information for a “localized” signal region in time. It also does not show how frequency information is changing from the start of the signal to its end, or at a particular signal region. Hence, it is mainly useful for stationary signals and pseudo-stationary signals, where the frequency information does not change during the signal acquisition.

### **1.1.2 Effects of Noise Presence on Signal Information in Frequency Domain**

The presence of random noise (mainly Gaussian) generated from instruments reduces the resolution in the signal domain, increasing the uncertainty in the information needed to reliably carry out a particular study. Because white random Gaussian noise contains all frequencies, there is overlap between signal and noise frequencies that cannot be distinguished in the frequency domain. In addition, there is a region in the frequency domain where only noise frequency exists. The greater the noise level in the signal domain, the more prominent is the energy density of the noise frequencies, further reducing the signal information. *Therefore, FT-based approaches fundamentally lack the capability to remove or reduce noise from the signal for the frequency range that contains the signal.*

Fig. 1.3 shows that frequency spectrum of model/synthetic cw-ESR and pulsed dipolar signals, respectively, with and without the presence of noise. The subfigure b1 reveals

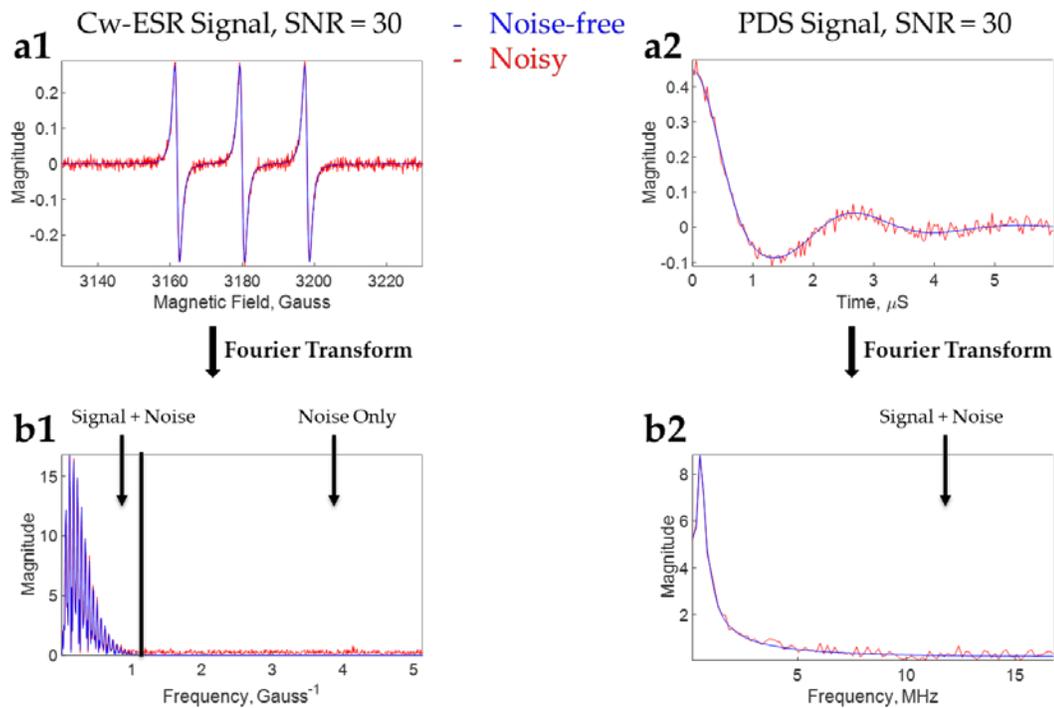
that the noise frequencies are present both at signal frequencies as well as at frequency regions where there is no signal frequency. The situation is especially true when the signal is only present at some frequency region such as low frequencies, as seen in the cw-ESR signal (Fig 1.3-b1). On the other hand, the subfigure b2 reveals a case when signal and noise frequencies are present at all frequency ranges. The effect of noise depends on their strength relative to that of the signal frequencies. In the case of a pulsed dipolar signal, distortion caused due to the noise increases at higher frequencies as signal presence decreases. With the signal dominant at lower frequencies, the effects of noise is limited.

It can be concluded from these two cases that removing a frequency where signal and noise both are present would result in signal distortion. Whereas, just removing the frequencies at the noise-only region would still contain noise at other frequencies.

The noise level is measured in terms of Signal-to-Noise Ratio (*SNR*) that is defined as,

$$SNR = \frac{Signal_{peak}}{rms(Noise)} \tag{1.3}$$

where *rms* is the root mean square of the noise levels at all the signal locations.



**Figure 1.3:** Effects of noise in the frequency domain in the cw-ESR and pulsed dipolar spectroscopy (PDS) signal. **a1)** Model cw-ESR signal with and without noise, **a2)** PDS signal with and without noise, **b1)** Fourier-transformed noise-free and noisy cw-ESR signal, and **b2)** Fourier-transformed noise-free and noisy PDS signal. The noisy signal has  $SNR = 30$ , measured using equation 1.3. Noise-free signals and spectra are in blue color and noisy signals and spectra are in red color.

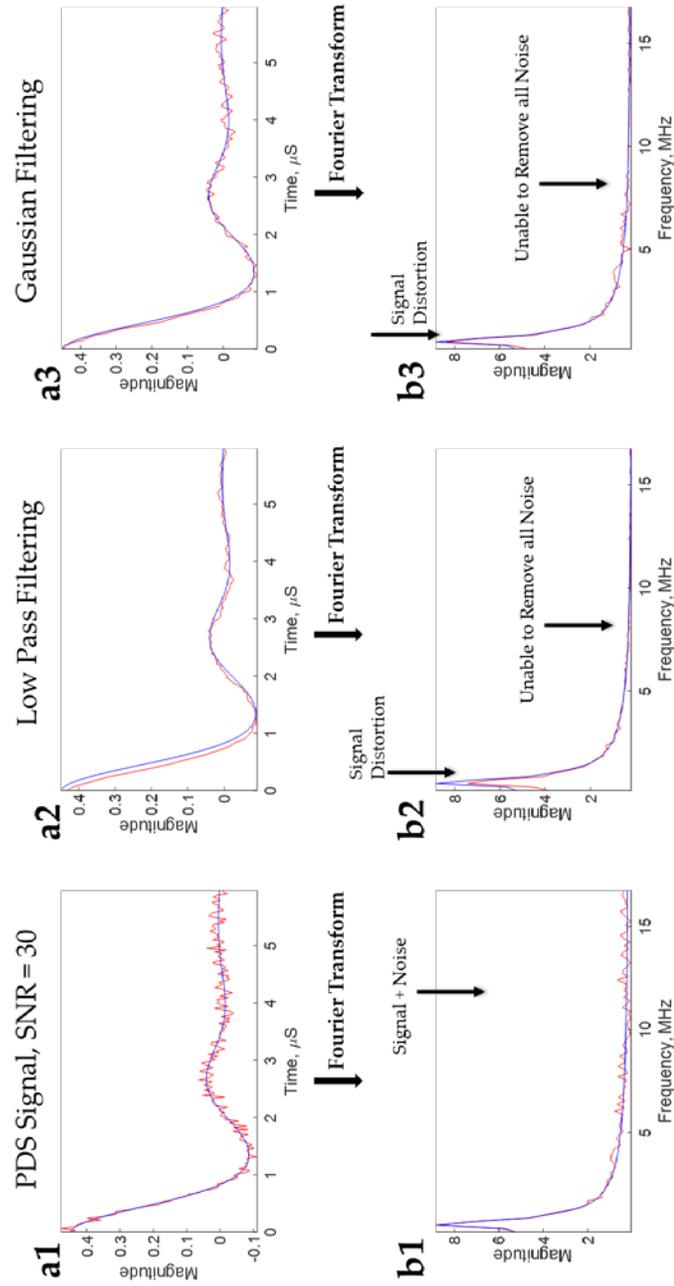
### 1.1.3 Frequency-based Noise Filtering and its Limitations

Despite the inability of FT in separating noise frequencies from signal frequencies, many filtering methods have been developed that use frequency information to reduce noise. A common method is to use low pass filtering, high pass filtering or band pass filtering. These methods retain certain frequency ranges, while eliminating the remaining frequencies. The selection of the frequency range is either based on *a priori* signal information or a reasonable assumption on frequency cutoff that is boundary between overlapping signal and noise frequencies and noise-only frequencies. However, these filtering methods do not eliminate noise frequencies that overlap with signal, leading to incomplete noise removal. Moreover, they are vulnerable to poor frequency cut-off that can lead to removal of signal frequencies, resulting in signal distortion. Many experimental signals are dominated by low frequencies. Applying low pass filtering do not remove low frequency noise and distort the signal information by removing the high frequency components of the signal. They end up smoothing the signal at the cost of resolution.

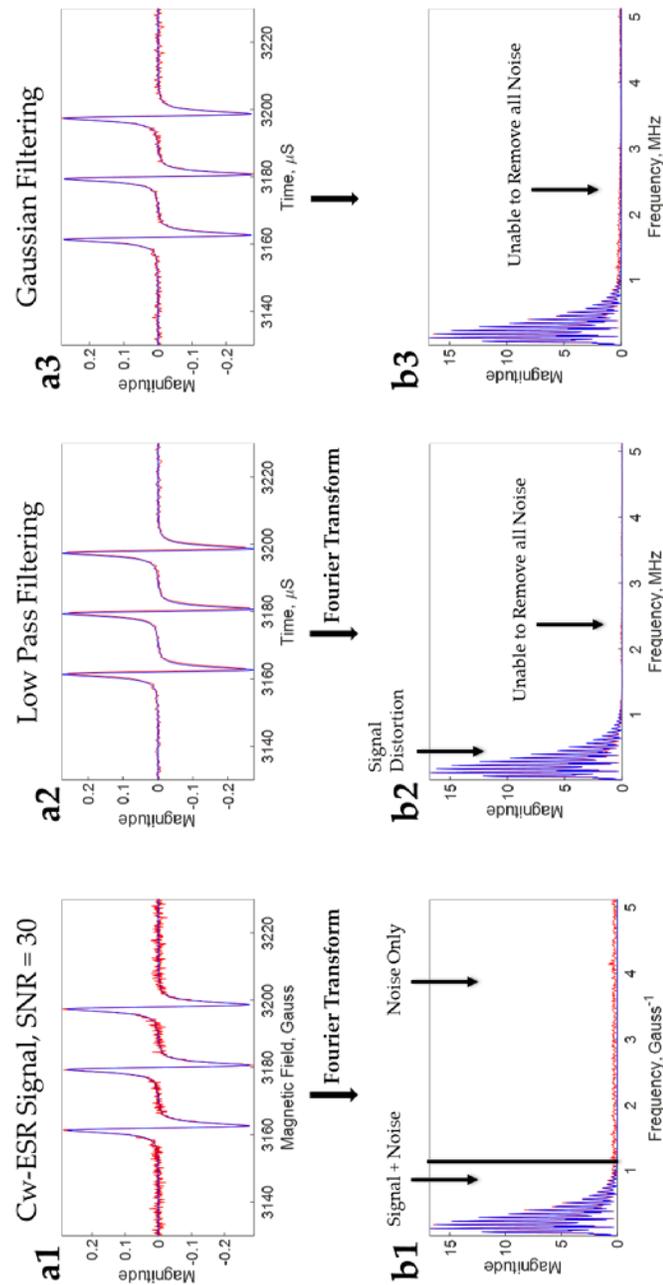
Figs. 1.4-1.5 show the frequency cut-off for low pass filtering method for the cw-ESR and pulsed dipolar signals, respectively. The low pass filter was used because these signals are low frequency signals as shown in Figs. 1.1 and 1.2.

Another filtering approach is to design a filter based on known signal or noise characteristics. These include linear filters (such as mean filter, Gaussian filter, Binomial filter, Weiner filter, Hamming window filter and anisotropic filter) and non-

linear filters (such as median and weighted-median filters). In addition to only retaining frequencies where signal exists, each signal frequency is also weighted as some frequencies are more important in representing the signal than others. They are also the frequencies with higher magnitude (i.e. energy density). This filtering approach (using Gaussian filter) performs better than low pass filtering (and also others), as shown in Figs. 1.4-1.5; however, it still lacks the capability to effectively remove the noise and is vulnerable to signal distortion. This is because overlapping noise and signal frequencies cannot be separated in the frequency domain, and filter design is imperfect in removing all the noise frequencies.



**Figure 1.4:** Result of low pass and Gaussian filtering methods for pulsed dipolar signal. **a1)** Model and noisy signal ( $SNR = 30$ ); **a2)** Model and noisy signal after low pass filtering; **a3)** Model and noisy signal after Gaussian filtering; **b1)** Frequency information of model and noisy signal; **b2)** Frequency information of model and noisy signal after low pass filtering; **b3)** Frequency information of model and noisy signal after Gaussian filtering. Model signal and its frequency information is in blue, whereas noisy and filtered signals and their frequency information are in blue.



**Figure 1.5:** Result of low pass and Gaussian filtering methods for cw-ESR signal. **a1)** Model and noisy signal ( $SNR = 30$ ); **a2)** Model and noisy signal after low pass filtering; **a3)** Model and noisy signal after Gaussian filtering; **b1)** Frequency information of model and noisy signal; **b2)** Frequency information of model and noisy signal after low pass filtering; **b3)** Frequency information of model and noisy signal after Gaussian filtering. Model signal and its frequency information is in blue, whereas noisy and filtered signals and their frequency information are in blue.

## 1.2 Short Time Fourier Transform

A modified version of FT, the Short Time Fourier Transform (STFT)<sup>1</sup>, can help to localize the signal to obtain a signal domain-frequency representation for analysis. A window is introduced in STFT to divide the signal into small time intervals and a FT of each segment is then performed, providing the frequency information at each interval. The SFFT of a signal can be obtained by<sup>3</sup>,

$$F(\tau, \omega) = \int_{-\infty}^{+\infty} f(t) w(t - \tau) e^{-i\omega t} dt \quad (1.4)$$

where  $f(t)$  is the input signal,  $w(t - \tau)$  is the rectangular window function, and  $F(\tau, \omega)$  is the FT of the signal lying within the window function.  $\tau$  is the translation parameter for sliding the window function through the signal.

The original signal can be reconstructed using the inverse STFT as given below,

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(\tau, \omega) e^{i\omega t} d\tau d\omega \quad (1.5)$$

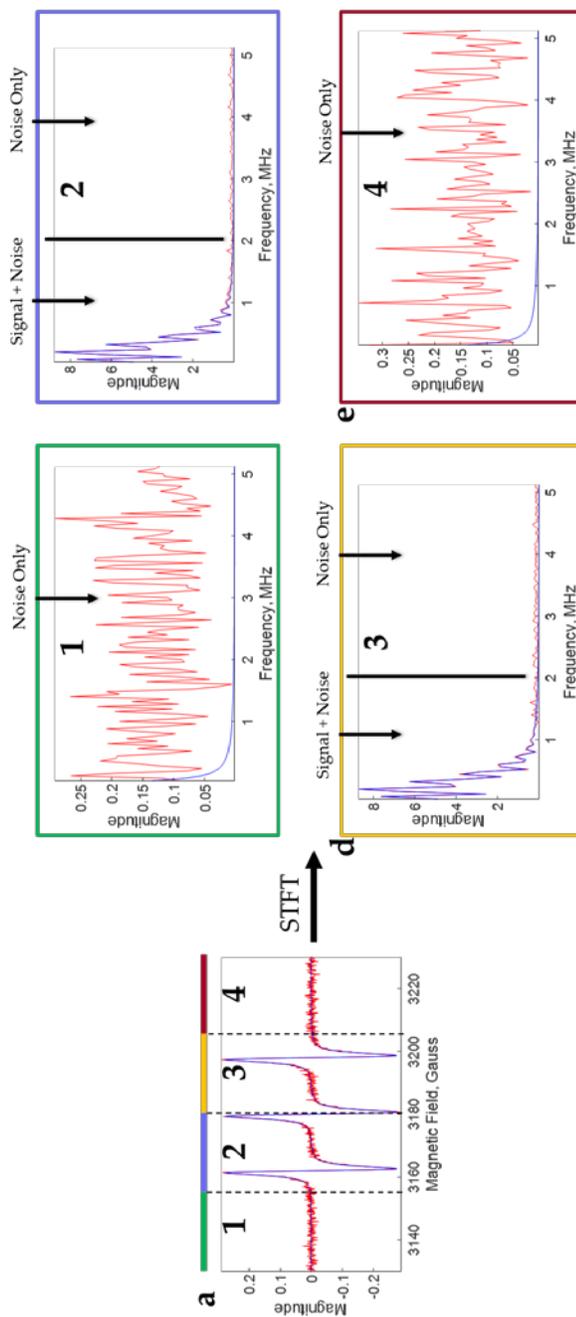
### 1.2.1 Limitations of STFT

However, there are conceptual limitations and practical challenges in implementing the STFT. The STFT assumes that the non-stationary signal is (approximately)

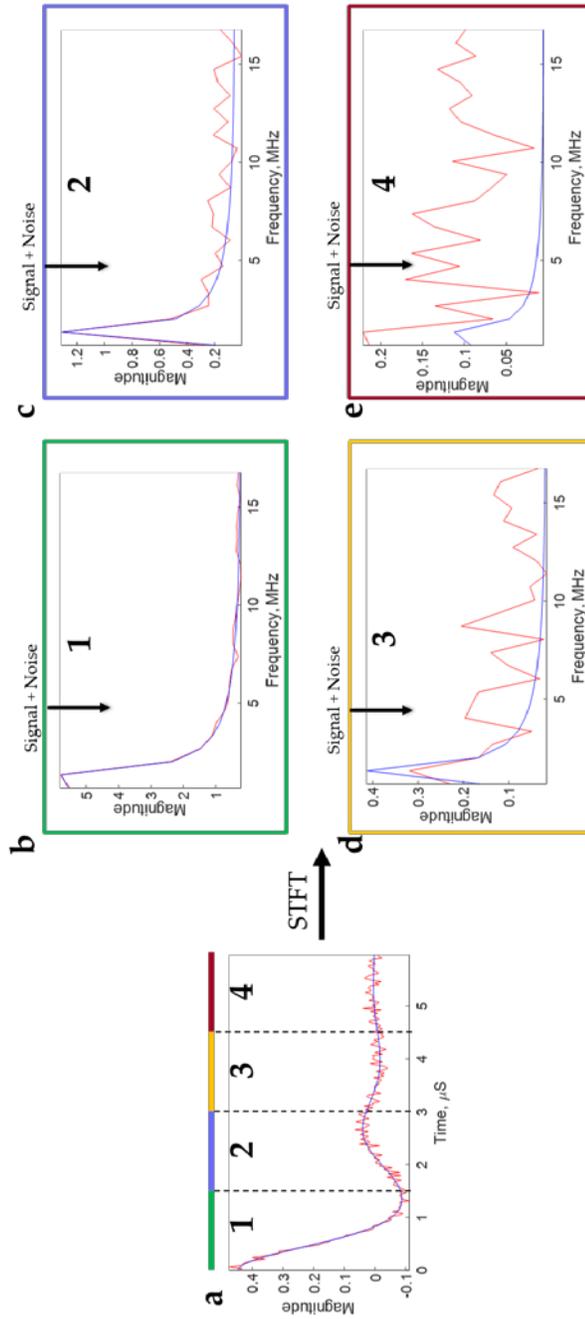
stationary in the finite intervals (as needed for FT), which is unlikely the case with real signals. The windows are finite in length, limiting the resolution that can be obtained in the frequency domain. The frequency resolution reduces with smaller windows.

The practical challenge entails the selection of an appropriate window size which is not well defined and is usually selected arbitrarily. Another challenge is the use of the translation parameter,  $\tau$ . Employing a non-overlapping window function on the signal creates a discontinuity between the adjacent signal intervals as well as loses the frequency information among the neighboring signal points in the adjacent intervals. Whereas selection of overlapping windows generates several Fourier-transformed spectra containing redundant information. It also becomes difficult to analyze several spectra. For instance, the STFTs of signal length 200 and window length 5 with translation of 1 would result in 204 ( $200+5-1$ ) frequency spectra. The length of experimental pulsed dipolar signal is usually between 200 and 300, whereas a cw-ESR signal usually has a length of 4096. The STFT in both cases will generate a very large number of FTs, making the study challenging.

Figs. 1.6 and 1.7 show STFT spectra for cw-ESR and pulsed dipolar signal using non-overlapping STFT. For easy comprehension, window length of 256 and 50 was selected to ensure only 4 FT spectra.



**Figure 1.6:** Short Time Fourier Transform (STFT) of model cw-ESR signal with and without noise, both of which are divided into 4 equal parts shown in green, purple, yellow and brown. **a)** Model cw-ESR with and without noise (Noisy Signal  $SNR = 30$ ), **b)** STFT of the 1<sup>st</sup> part (first 1/4<sup>th</sup>) of the signal in green, **c)** STFT of the 2<sup>nd</sup> part (second 1/4<sup>th</sup>) of the signal in purple, **d)** STFT of the 3<sup>rd</sup> part (third 1/4<sup>th</sup>) of the signal in yellow, and **e)** STFT of the 4<sup>th</sup> part (fourth 1/4<sup>th</sup>) of the signal in brown. Model signal without noise is in blue color, whereas model signal with noise is in red color.



**Figure 1.7:** Short Time Fourier Transform (STFT) of model pulsed dipolar signal with and without noise, both of which are divided into 4 equal parts shown in green, purple, yellow and brown. **a)** Model signal with and without noise (Noisy Signal  $SNR = 30$ ), **b)** STFT of the 1<sup>st</sup> part (first  $1/4^{\text{th}}$ ) of the signal in green, **c)** STFT of the 2<sup>nd</sup> part (second  $1/4^{\text{th}}$ ) of the signal in purple, **d)** STFT of the 3<sup>rd</sup> part (third  $1/4^{\text{th}}$ ) of the signal in yellow, and **e)** STFT of the 4<sup>th</sup> part (fourth  $1/4^{\text{th}}$ ) of the signal in brown. Model signal without noise is in blue color, whereas model signal with noise is in red color.

### 1.2.2 Limitations of STFT-based Noise Filtering

Although the STFT-based filtering methods, in principle, can perform better than the FT method, they are rarely used because of its implementation challenges and marginally improved results compared to the FT-based filtering methods. More importantly, the STFT, like the FT, cannot separate overlapping noise and signal frequencies within the signal interval selected by window function. The method is also prone to generating artifacts during noise filtering.

### 1.3 Wavelet Transform

A wavelet transform (WT) can simultaneously represent signal domain-frequency information for analysis through signal localization and is defined as<sup>1,3,4</sup>,

$$F(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{+\infty} f(t) \psi^* \left( \frac{t - \tau}{s} \right) dt \quad (1.6)$$

where  $s$  is the inverse frequency (or frequency range) observing parameter (also called scale parameter),  $\tau$  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 “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 complex conjugate of  $\psi(t)$ , which for a real function is the same ( $\psi^*(t) = \psi(t)$ ).

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

Unlike STFT, the WT employs variable window width and 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 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 wavelet domain is spatially correlated with the location of 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,

$$f(t) = \frac{1}{C_\psi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{s^2} F(\tau, s) \psi\left(\frac{t-\tau}{s}\right) d\tau ds \quad (1.7)$$

where  $C_\psi$  is admissibility constant which is written as,

$$C_\psi = \sqrt{2\pi \int_{-\infty}^{+\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega} < \infty \quad (1.8)$$

where  $\Psi(\omega)$  is the FT of the wavelet function  $\psi(t)$ . The constraint in equation 1.8 implies that the wavelet function  $\psi(t)$  must also be oscillatory with zero mean i.e.,  $\int_{-\infty}^{+\infty} \psi(t) dt = 0$ .

### 1.3.1 Heisenberg Uncertainty Boxes for Signal Domain-Frequency Resolution

There is fundamental limit in achieving simultaneous signal domain and frequency resolution using wavelets. It is called "Heisenberg Boxes"<sup>1</sup> based on the uncertainty principle that is defined as:  $\sigma_t \sigma_f \geq \frac{1}{4\pi}$ , where  $\sigma_t$  is the standard deviation of the wavelet function in time and  $\sigma_f$  the standard deviation of the wavelet function in frequency. Both of these represent the uncertainty in resolving time and frequency information of a signal simultaneously. The more accurate the time resolution, the less precise is the frequency resolution and vice-versa. This also implies that the scaling

parameter  $s$  representing frequency and translation parameter  $\tau$  representing time are not completely independent variables, and must conform to the uncertainty principle in the  $s, \tau$  plane.<sup>3</sup>

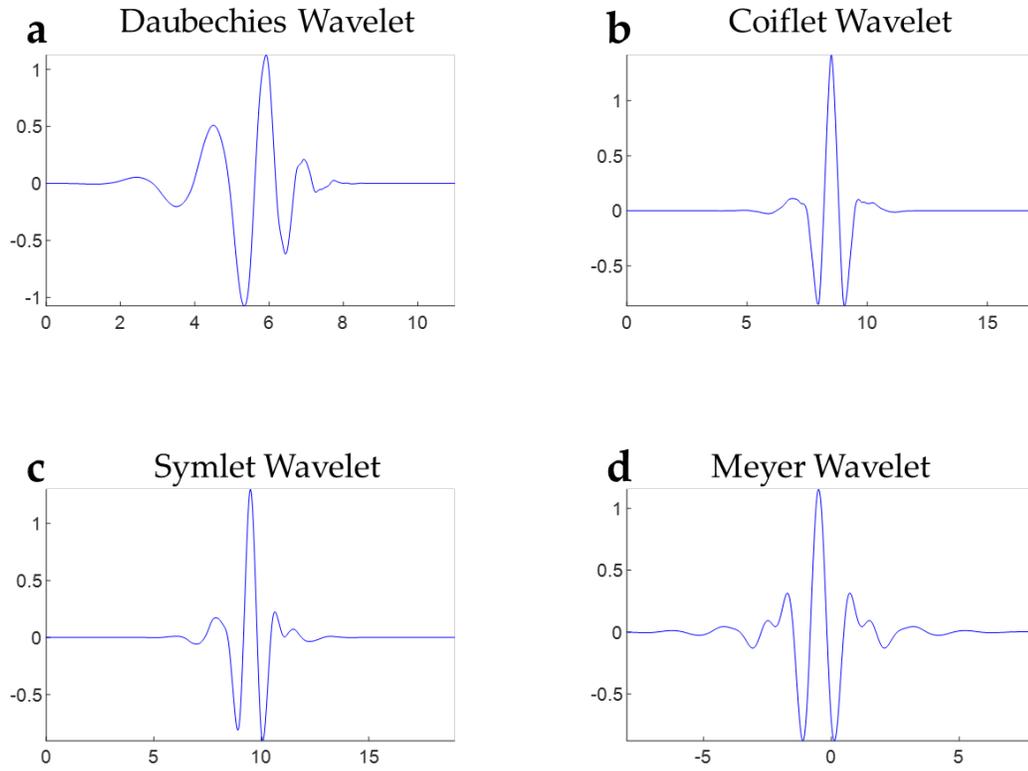
### 1.3.2 Wavelet Selection

Unlike FT and STFT, the WT can use a variety of wavelet functions for signal analysis. Different wavelets are used to vary selectivity or sensitivity of adjacent frequencies with respect to signal localization, in accordance with Heisenberg boxes. A wavelet can have either a different shape of uncertainty boxes within signal domain–frequency plane or can have a different shape compared to other wavelets. The uncertainty in signal and frequency resolution does not depend on a *priori* information of the signal or its characteristics, but on the wavelet characteristics. Depending on the application, different wavelets can be selected. For removing noise from signals, the distinct properties of wavelets can permit much improved separation and identification of noise from signal because signal and noise typically have distinct signal domain-frequency characteristics.

There are many standard wavelet families such as Daubechies, Coiflets, Symlets, and Meyer, among others, that can be used to analyze any signal (cf. Fig. 1.8). Within each wavelet family, there are many wavelet functions that represent the same wavelet properties. They differ in function length and resolution capability. In addition, there is a special wavelet called Haar, which is designed from the rectangular function. With

window length variable, it is the generalized form of the rectangular window used in STFT. A customized wavelet can also be developed for signals whose features are well-known and predetermined. An example of such wavelets are CDF 5/3 and CDF 9/7 (Cohen-Daubechies-Feauveau) that are used in image processing applications.

In general, any wavelet can be used to effectively analyze a signal. A good choice of wavelet results in a better analysis compared to other wavelets for a particular application. For instance, the minimum SNR required in order to remove noise from signal can vary by the choice of wavelet, depending upon how well a wavelet can resolve their distinct features. The best wavelet to use for the signal will require a minimum SNR needed to successfully retrieve signal. For cw-ESR and pulsed dipolar signals, Coiflet and Daubechies, respectively, performed better than other wavelets families.



**Figure 1.8:** Wavelet function of some standard wavelet families. **a)** Daubechies wavelet (daubechies-6), **b)** Coiflet wavelet (coiflet-3), **c)** Symlet wavelet (symlet-10), **d)** Meyer wavelet. “daubechies-6”, “coiflet-3” and “symlet-10” are specific wavelet functions within the wavelet family.

## 1.4 Discrete Wavelet Transform

The WT in equation 1.6 is impractical due to the high computational complexity and high resource requirements. Also, it results in several wavelet-transformed signals that contain redundant information and are impractical to analyze (like STFT). Similar to the Fast Fourier Transform (FFT), the discrete wavelet transforms (DWT) has been developed to obtain the WT at a reduced computation complexity (to the order of  $\mathcal{O}(p)$ , where  $p$  is the length of input signal) with a minimum number of wavelet coefficients required to represent non-overlapping signal information. The DWT of a discrete input signal is defined as,

$$D[a, b] = \frac{1}{\sqrt{b}} \sum_{m=0}^{p-1} f[t_m] \psi \left[ \frac{t_m - a}{b} \right] \quad (1.9)$$

where,  $D[a, b]$  is the wavelet-transformed signal,  $f[t_m]$  is the discrete input signal, the translation parameter  $\tau = a$  and the scale parameter  $s = b$  ( $a$  and  $b$  both being integers).

However, for practical purposes, the translation and scale parameters are discretized as  $a = k2^{-j}$  and  $b = 2^{-j}$ , where  $j$  and  $k$  are integers referring to scale index and wavelet-transformed signal index, respectively. The wavelet function in equation 1.9 can be rewritten as,

$$\psi_{j,k}[t_m] = \frac{1}{\sqrt{2^{-j}}} \psi \left[ \frac{t_m - k2^{-j}}{2^{-j}} \right] \quad (1.10)$$

or,

$$\psi_{j,k}[t_m] = 2^{j/2} \psi[2^j t_m - k] \quad (1.11)$$

This power of two discretization of WT is called dyadic decomposition. It is both computationally most efficient and practically simple due to binary implementation in machines. Also, this dyadic decomposition divides the frequency information of the signal into non-overlapping frequency sub-bands, eliminating the redundancy due to repetition of signal domain-frequency information in the wavelet domain. To represent non-overlapping frequency sub-bands, the dyadic wavelet needs to be orthonormal satisfying the following condition,

$$\sum_{t_m=0}^{L-1} \psi_{j,k}[t_m] \psi_{\tilde{j},\tilde{k}}[t_m] = \begin{cases} 1 & \text{if } j = \tilde{j} \text{ and } k = \tilde{k} \\ 0 & \text{otherwise} \end{cases} \quad (1.12)$$

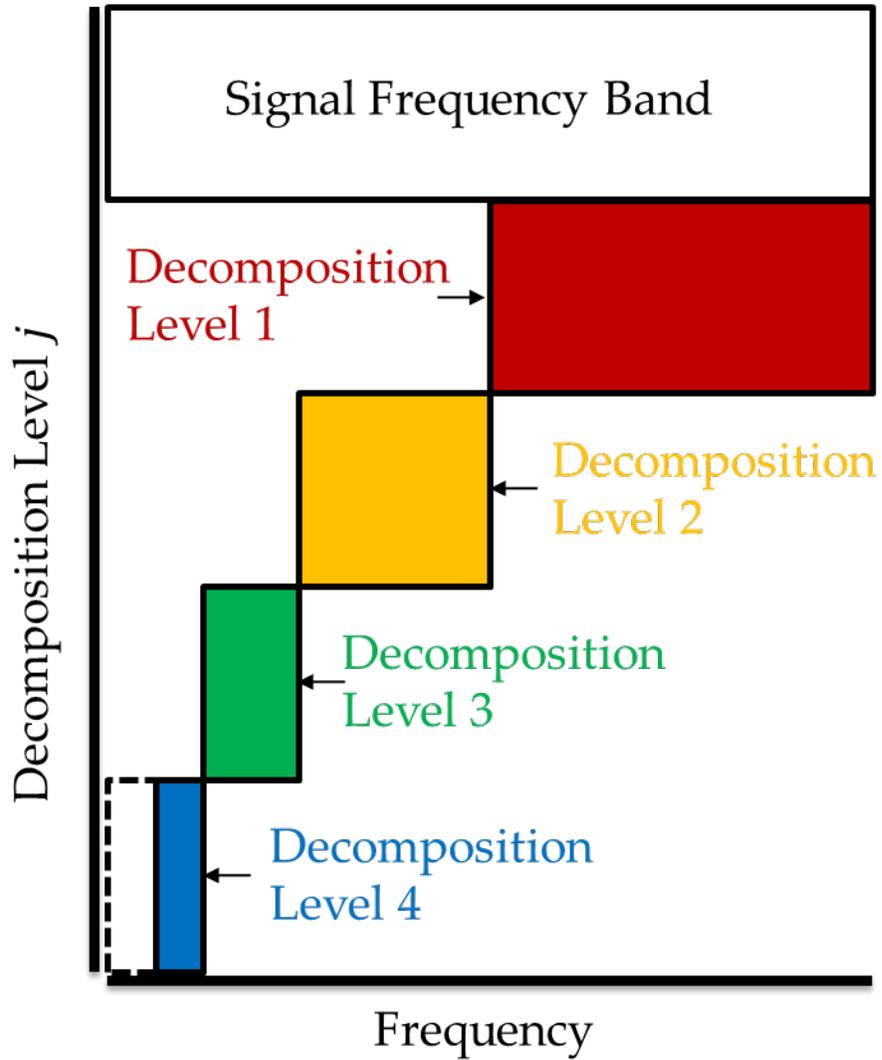
Incorporating equation 1.11 into equation 1.9, the DWT can be written as,

$$D_j[k] = \sum_{m=0}^{p-1} f[t_m] \psi_{j,k}[t_m] \quad (1.13)$$

where  $D_j[k]$  is the wavelet-transformed signal at scale  $j$ . The scale  $j$  represents a frequency sub-band of the overall signal. From equation 1.11, note that the frequency

sub-band representation moves from higher frequencies to lower frequencies with increase in scale from  $j$  to  $j + 1$ . It is due to the dilation (i.e., expansion) of the wavelet function that increases the signal coverage by the wavelet, and hence reduces the frequency sub-band observed (cf. Fig 1.9). This approach is referred as sub-band decomposition, and because the scale  $j$  is associated with it, it is called decomposition at  $j^{th}$  level or  $j^{th}$  decomposition level. Furthermore, the wavelet-transformed signal  $D_j[k]$  is also called the “Detail component”. This is because  $D_j[k]$  represents non-zero frequency which represents higher frequencies observing change in signal from one location to other. In other words,  $D_j[k]$  informs about the details in the signal represented by higher frequencies. That is why it is always centered around zero because it does not contain any zero frequency component.

The decomposition level  $j$  cannot be infinite for a finite input signal  $f[t_m]$  with length  $p$ . The dyadic decomposition allows up to  $N$  Detail components where  $N = \lfloor \log_2 p \rfloor$ . Therefore, the decomposition level  $j$  can range  $1 \leq j \leq N$ .



**Figure 1.9:** The frequency sub-band coverage for Detail component from decomposition levels 1 to 4. Remaining decomposition levels can be constructed from the dotted frequency sub-band adjacent to the decomposition level 4. Each decomposition level represent the non-overlapping frequencies of the input signal.

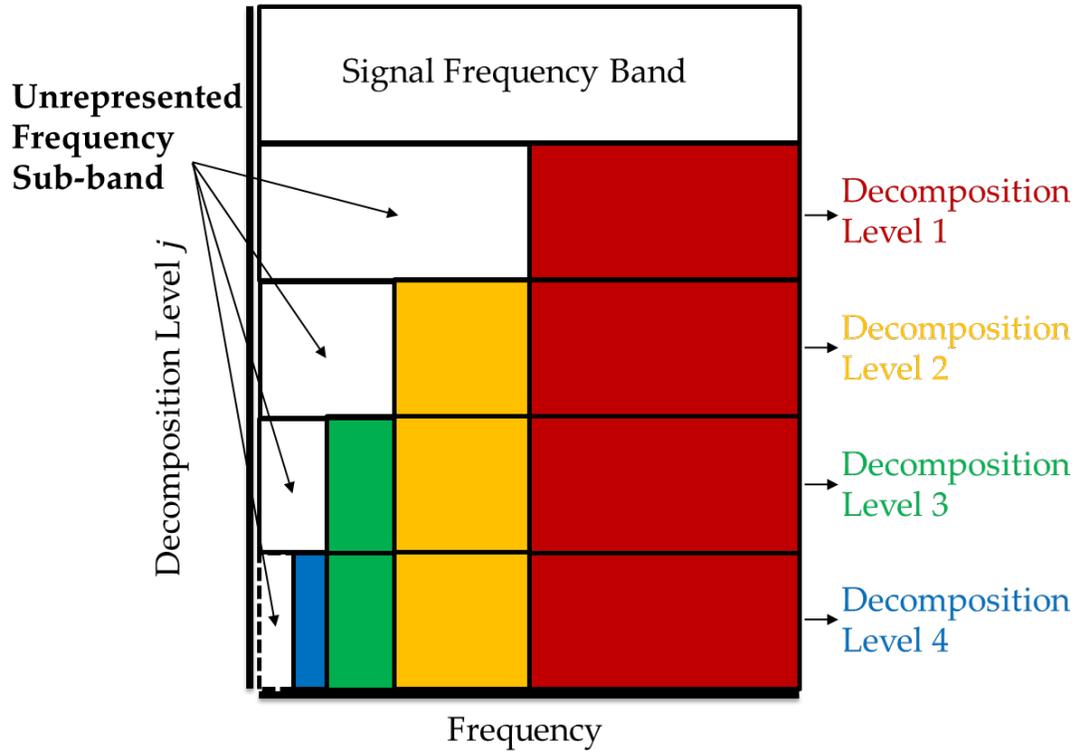
### 1.4.1 Introduction to Approximation Component

The collection of all frequency sub-bands represented by the Detail components does not constitute complete frequency band of the input signal. They only cover the frequency sub-bands represented till the  $j^{th}$  decomposition level. Even with the largest decomposition level  $N$ , some low frequencies remain unrepresented. This phenomenon is due to the discrete nature of the signal and the DWT. Fig. 1.10 shows the unrepresented frequency region at the  $j^{th}$  decomposition level.

To represent the remaining low frequencies, an Approximation component at the  $j^{th}$  decomposition level ( $1 \leq j \leq N$ ) is obtained. It is acquired as follows,

$$A_j[k] = \sum_{m=0}^{p-1} f[t_m] \phi_{j,k}[t_m] \tag{1.14}$$

where  $A_j[k]$  is the Approximation component and  $\phi_{j,k}[t_m]$  is the scaling function.  $A_j[k]$  is called the Approximation component because it represents the low frequency information including the zero frequency component that constitutes the low resolution approximation of a signal. Note that the discrete input signal  $f[t_m]$ , in principle, may be thought of as an Approximation component at the  $0^{th}$  decomposition level of  $f(t)$ , i.e.  $s = 1$  and  $\tau = k$ .



**Figure 1.10:** The frequency sub-band showing unrepresented frequencies after coverage of Detail components from decomposition levels 1 to 4.

### 1.4.2 Scaling Function and its relation to Wavelet Function

The scaling function (also called the father wavelet),  $\phi_{j,k}[t_m]$ , is designed to represent frequencies not covered by the wavelet function. Similar to the wavelet function in equation 1.11, it is written as follows,

$$\phi_{j,k}[t_m] = 2^{j/2} \phi[2^j t_m - k] \quad (1.15)$$

As scaling and wavelet function represent non-overlapping frequency sub-bands of the input signal at a decomposition level, they are orthogonal to each other and must satisfy the following condition:

$$\sum_{t_m=0}^{L-1} \phi_{j,k}[t_m] \psi_{j,k}[t_m] = 0 \quad (1.16)$$

where  $L$  is the finite length of the scaling and wavelet function and their values are selected depending on the choice of wavelet. The scaling function is also orthogonal to all the wavelet functions from decomposition levels from 1 to  $j - 1$ , because they are the frequency sub-bands that have already been represented by Detail components containing higher frequencies. Therefore, to incorporate this orthogonality, equation 1.16 can be modified as,

$$\sum_{t_m=0}^{L-1} \phi_{j,k}[t_m] \psi_{i,k}[t_m] = 0 \quad (1.17)$$

where  $1 \leq i \leq j$ .

In other words, the scaling function is the aggregation of all other unobserved frequency sub-bands and can be defined in terms of the wavelet function by,

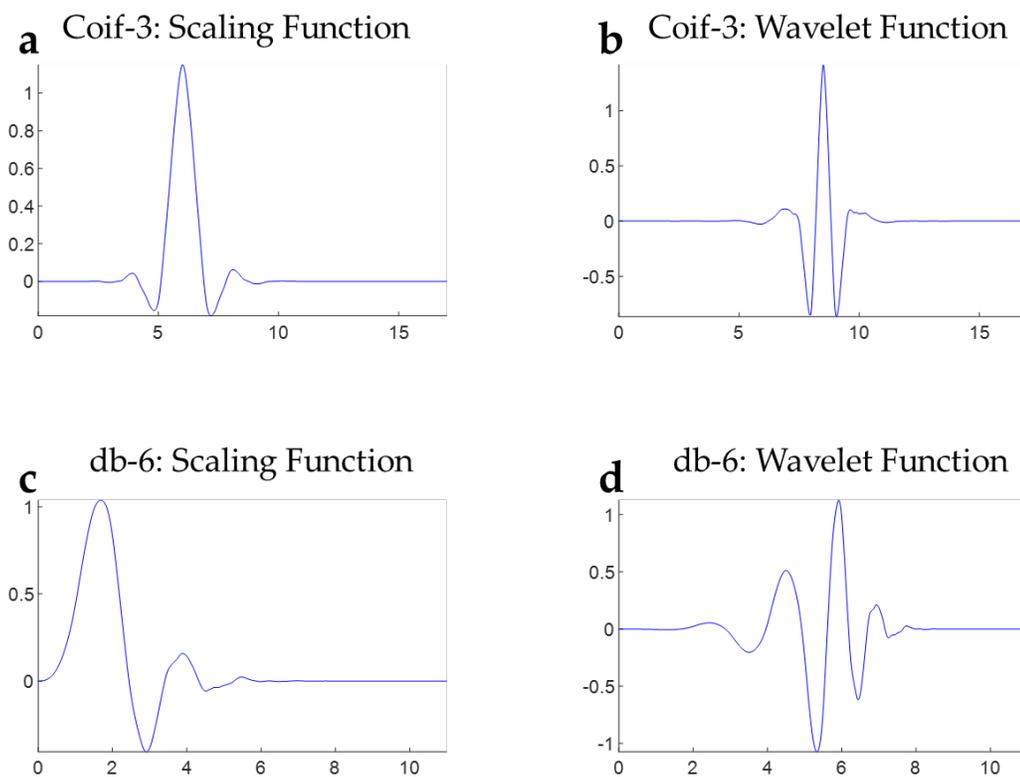
$$|\tilde{\phi}(\omega)|^2 = \sum_{i=j+1}^{\infty} \frac{|\Psi(2^{-i}\omega)|^2}{2^{-i}} \quad (1.18)$$

At the  $N^{th}$  decomposition level, the equation is,

$$|\tilde{\phi}(\omega)|^2 = \sum_{i=N+1}^{\infty} \frac{|\Psi(2^{-i}\omega)|^2}{2^{-i}} \quad (1.19)$$

where  $\tilde{\phi}(\omega)$  is the FT of  $\phi(t_m)$ .

In Fig 1.11, the scaling and wavelet function of the two wavelets (“coif-3” and “db-6”) are shown. These two wavelets are used for denoising cw-ESR (“coif-3”) and pulsed ESR (“db-6”) signals.



**Figure 1.11:** Scaling and Wavelet functions of Coiflet-3 (Coif-3) and Daubechies-6 (db-6) wavelets belonging to Coiflet and Daubechies wavelet families. **a)** scaling function of Coiflet-3 wavelet, **b)** wavelet function of Coiflet-3 wavelet, **c)** scaling function of Daubechies-6 wavelet, and **d)** wavelet function of Daubechies-6 wavelet. These wavelets are used for cw-ESR (Coif-3) and pulsed dipolar signal (db-6).

### 1.4.3 Inverse Discrete Wavelet Transform

The input signal  $f[t_m]$  can be reconstructed using inverse discrete wavelet transform (IDWT) in the following way:

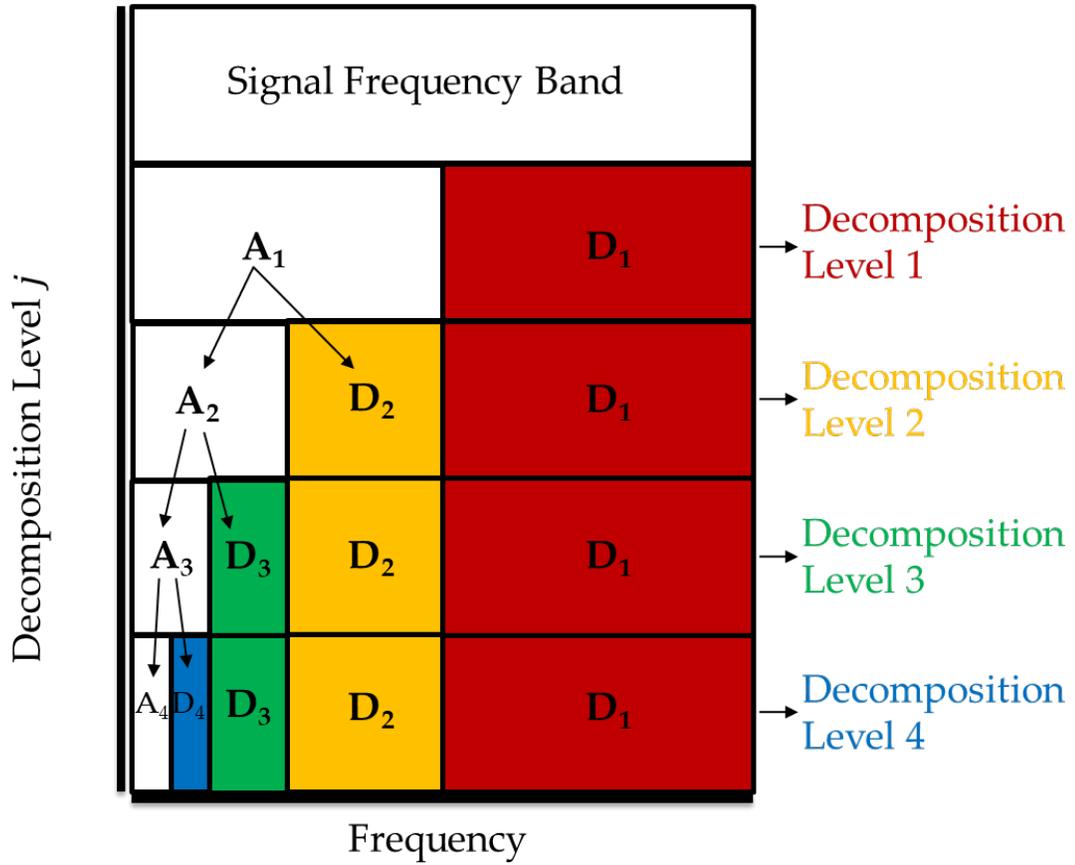
$$f[t_m] = \sum_{k=0}^{p-1} A_N[k] \phi_{N,k}[t_m] + \sum_{j=1}^N \sum_{k=0}^{p-1} D_j[k] \psi_{j,k}[t_m] \quad (1.20)$$

Further, at a given decomposition level  $j_0$ , IDWT can be given as,

$$f[t_m] = \sum_{k=0}^{p-1} A_{j_0}[k] \phi_{N,k}[t_m] + \sum_{j=1}^{j_0} \sum_{k=0}^{p-1} D_j[k] \psi_{j,k}[t_m] \quad (1.21)$$

### 1.4.4 Relationship between $j^{th}$ Approximation Component, and $j + 1^{th}$ Approximation and Detail Component

The Approximation component at the  $j^{th}$  decomposition level contains frequency information that is not collectively represented by the Detail components from 1 to  $j^{th}$  decomposition levels. Therefore, for the  $j + 1^{th}$  decomposition level, the frequency sub-band of the Approximation component at  $j^{th}$  decomposition level is divided into two equal non-overlapping sub-bands of low and high frequency bands to obtain the  $j + 1^{th}$  Approximation and Detail components. Fig. 1.12 demonstrates the construction of the Approximation and Detail components at different decomposition levels.



**Figure 1.12:** The frequency sub-band coverage of the Approximation and Detail components up to decomposition level 4.

## 1.5 Fast Wavelet Transform

In the DWT, the Approximation and Detail component at each decomposition level is obtained from the input signal using the scaling and wavelet function, respectively. For fast computation, just like the use of the FFT in order to FT, the Fast Wavelet Transform (FWT) is employed that eliminates the need for calculating the scaling and wavelet function for each decomposition level. The input signal is only used to obtain the Approximation and Detail component at decomposition level 1. As described in section 1.4.4, the FWT utilizes the information that the frequency sub-band of the  $j^{th}$  Approximation component that can be equally divided into low and high frequency sub-bands, which also represent  $j + 1^{th}$  Approximation and Detail components, respectively. Therefore for the subsequent decomposition levels, say  $j + 1$ , the  $j^{th}$  Approximation component is used to acquire the  $j + 1^{th}$  Approximation and Detail components. The FWT uses low and high pass filters to replace scaling and wavelet functions, respectively, to obtain the Approximation and Detail components (cf. Fig. 1.14) as follows,

$$A_{j+1}[k] = \sum_{n=0}^{L-1} l[n] A_j[2k + n] \quad (1.22)$$

and,

$$D_{j+1}[k] = \sum_{n=0}^{L-1} h[n] A_j[2k + n] \quad (1.23)$$

where  $l[n]$  and  $h[n]$  are low and high pass signal decomposition filters, respectively, and  $L$  is the length of both the filters (and also the scaling and wavelet functions). The filter values are the coefficient values of the scaling and wavelet function.

As can be seen from equations 1.22 and 1.23, the low and high pass filters are recursively applied on the Approximation component to obtain the Approximation and Detail components, respectively, at the next level. For decomposition level 1, the input signal is the Approximation component at the  $0^{th}$  level. Fig. 1.13 shows the block diagram of the FWT process.

It must be noted in the figure that after applying low and high pass filters, the Approximation and Detail components are downsampled by 2, which means every alternate sample is removed from the output signal. In equation 1.22 and 1.23, it is reflected in  $2k + n$  term, where 2 represents every alternate data point. The downsampling preserves the total number of input and output data points in signal and wavelet domain, respectively. The signal can be perfectly reconstructed without any information loss because it follows the Nyquist rule.

The inverse Fast Wavelet Transform (IFWT) is defined as,

$$A_j[k] = \sum_{n=1}^{L-1} \tilde{l}[n] A_{j+1}[2k + n] + \sum_{n=1}^{L-1} \tilde{h}[n] D_{j+1}[2k + n] \quad (1.24)$$

where  $\tilde{l}[n]$  and  $\tilde{h}[n]$  are low and high pass signal reconstruction filters. Using equation 1.24, the input signal can be reconstructed as follows,

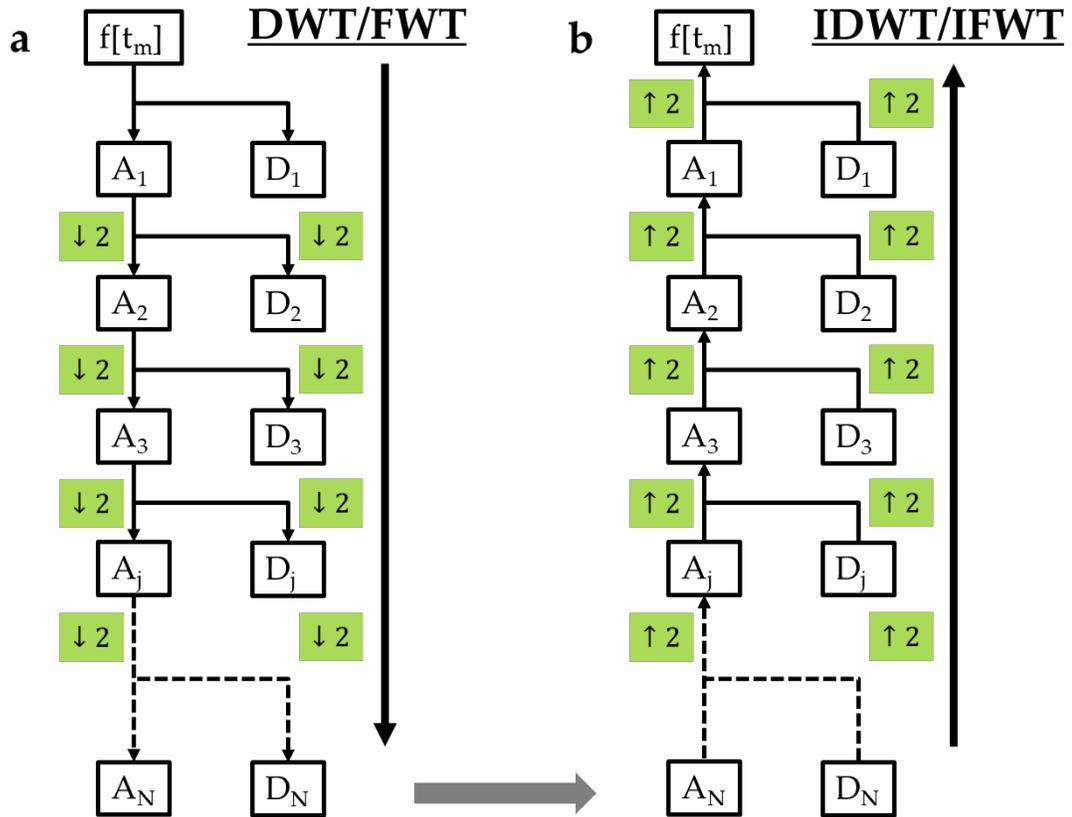
$$f[t_m] = \sum_{n=1}^{L-1} \tilde{l}[n] A_N[2k + n] + \sum_{j=1}^N \sum_{n=1}^{L-1} \tilde{h}[n] D_j[2k + n] \quad (1.25)$$

Further, at a given decomposition level  $j_0$ , the IFFT can be given as,

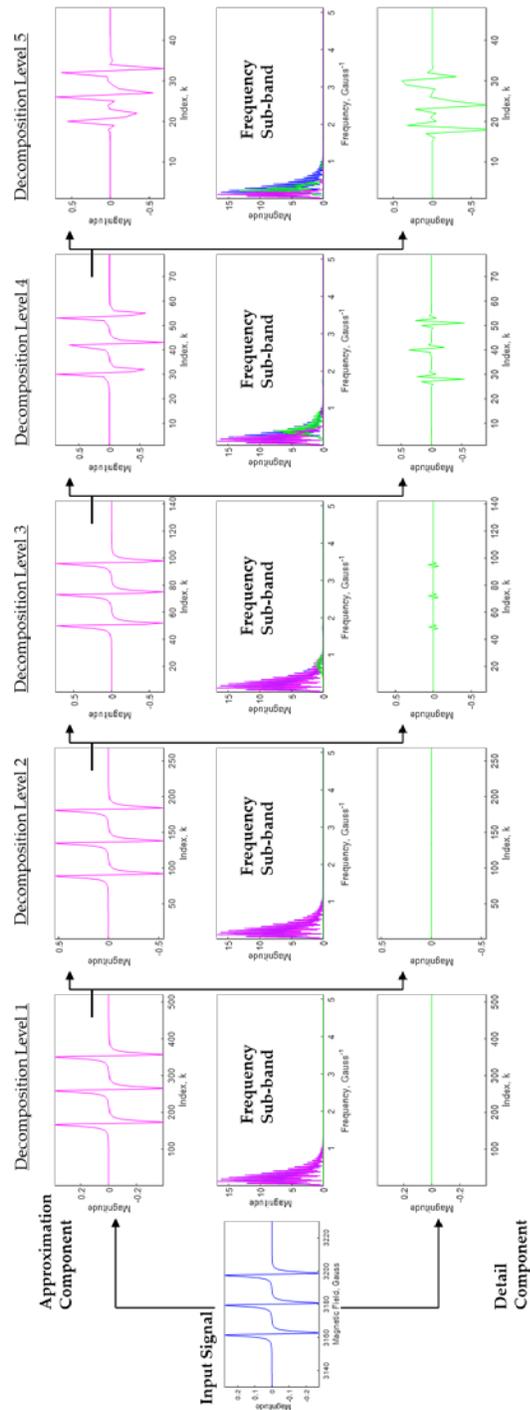
$$f[t_m] = \sum_{n=1}^{L-1} \tilde{l}[n] A_{j_0}[2k + n] + \sum_{j=1}^{j_0} \sum_{n=1}^{L-1} \tilde{h}[n] D_j[2k + n] \quad (1.26)$$

Equations 1.24 and 1.25 (also 1.26) reveal that low and high pass filters are applied on the  $j + 1^{th}$  Approximation and Detail components, respectively, to obtain the  $j^{th}$  Approximation component (cf. Fig. 1.15). This process can be applied recursively to reconstruct the input signal. Fig. 1.13 shows the block diagram of the IFFT process for input signal reconstruction. Converse to downsampling, it can be seen in the figure that the Approximation and Detail components are upsampled by 2 before applying the filters. The upsampling by 2 increases the signal length by 2 by inserting 0 between the two data points. The location of these data points are the same ones that were removed during the downsampling procedure.

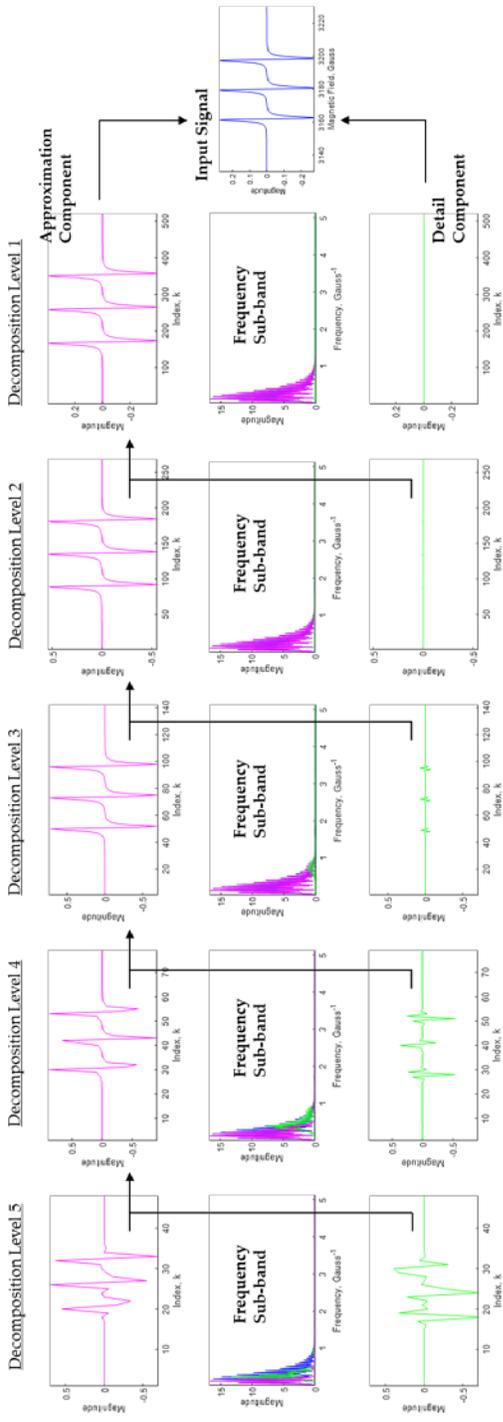
**Note:** The DWT and FWT are used interchangeable because FWT is the process to accomplish DWT in a computationally efficient manner.



**Figure 1.13:** The block diagram of **a)** Fast Wavelet Transform (FWT) and **b)** Inverse Fast Wavelet Transform (IFWT).  $f[t_m]$  is the input signal,  $A_j$  and  $D_j$  are the Approximation and Detail component, respectively, at  $j^{th}$  decomposition level obtained through recursive low pass filtering (for Approximation component) and high pass filtering (for Detail component), and  $\downarrow 2$  and  $\uparrow 2$  are downsampling and upsampling, respectively, by a factor of 2. The process of FWT and IFWT leads to Discrete Wavelet Transform (DWT) and Inverse Discrete Wavelet Transform (IDWT), and hence they are used interchangeable. Equations 1.22 and 1.23 describe the low pass and high pass filtering process.



**Figure 1.14:** The process of FWT illustrated with an example. The Approximation and Detail components at each decomposition level is displayed in magenta and green, respectively, along with their frequency-sub-band representation.



**Figure 1.15:** The process of IFWT illustrated with an example. The Approximation and Detail components at each decomposition-band is displayed in magenta and green, respectively, along with their frequency-sub-band representation.

### 1.5.1 Relationship between Low and High Pass Filters

As the scaling and wavelets functions are orthogonal to each other, the low and high pass filters are also designed to be orthogonal. For FWT, the relation between low and high pass filters is as follows,

$$h[n] = (-1)^n l[L - 1 - n] \quad (1.27)$$

Similarly, for the IFWT, the filters are related in the following way,

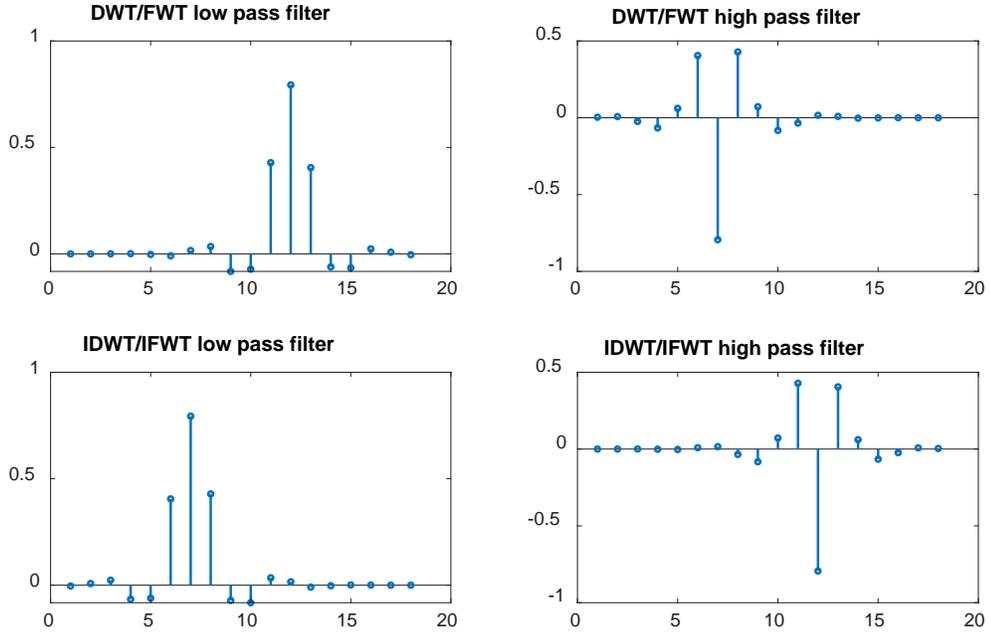
$$\tilde{h}[n] = (-1)^n \tilde{l}[L - 1 - n] \quad (1.28)$$

The relation between FWT and IFWT filters is as follows,

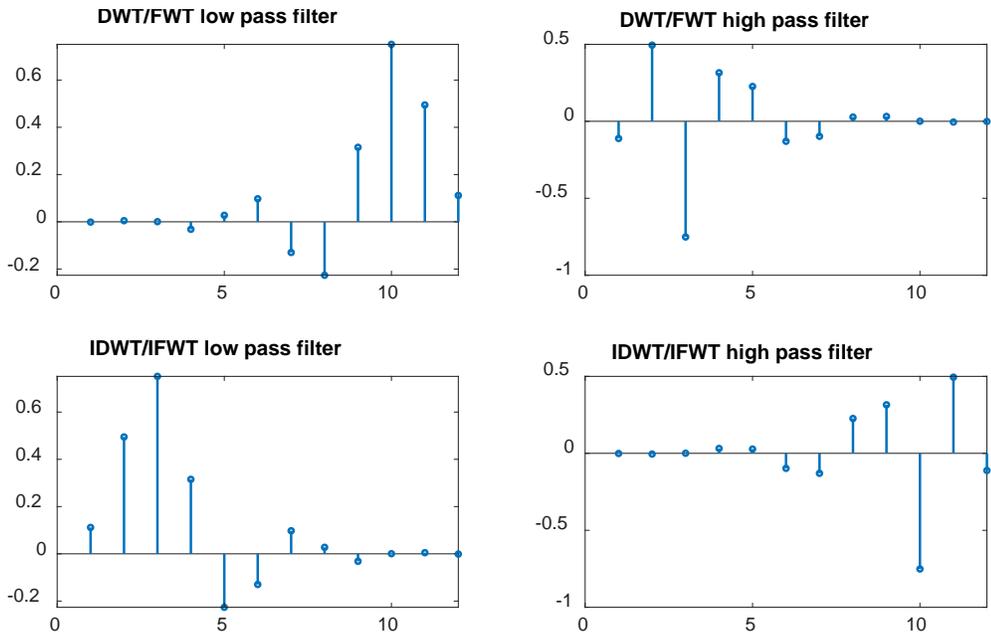
$$\tilde{h}[n] = h[L - 1 - n] \quad (1.29)$$

$$\tilde{l}[n] = l[L - 1 - n] \quad (1.30)$$

Figs. 1.16 and 1.17 show the low and high pass filters in the FWT and IFWT for coiflet-3 and Daubechies-6 wavelets, respectively. The values of the filters are listed in the Appendix 1. These wavelets are shown because they are used in cw-ESR and pulsed dipolar signal. In Appendix 1, the MATLAB script to generate the above filters is provided as well as information regarding obtaining filters for any wavelet is presented.



**Figure 1.16:** The plots of low and high pass filters of Coiflet-3 wavelet that are used in the FWT and IFWT for obtaining DWT and IDWT, respectively.



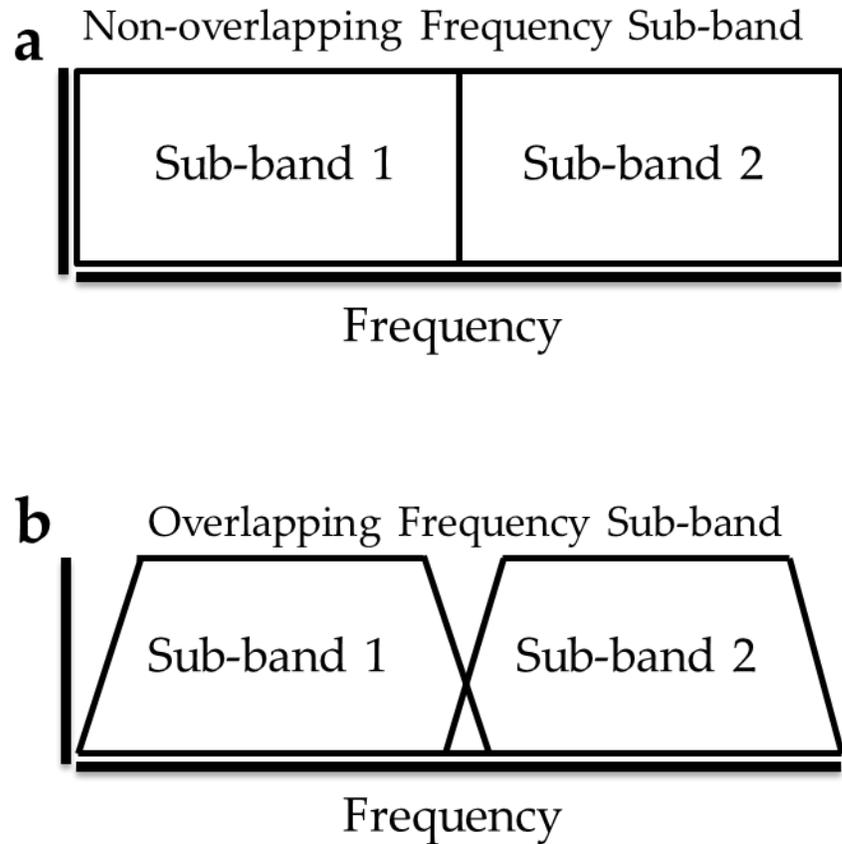
**Figure 1.17:** The plots of low and high pass filters of Daubechies-6 wavelet that are used in the FWT and IFWT for obtaining DWT and IDWT, respectively.

### 1.5.2 Orthogonal and Biorthogonal Wavelets

Although DWT and FWT are developed for wavelets that are orthogonal in principle (i.e., they represent non-overlapping information), in practice, both orthogonal and biorthogonal wavelets can be used. Biorthogonal wavelets are a set of wavelets that allows the wavelet transform to be invertible but may or may not be orthogonal. The biorthogonal wavelets are symmetric implying that the filters have linear phase. (Orthogonal wavelets, other than the Haar wavelet, are not symmetric). They use different filters at DWT and IDWT for perfect reconstruction of signals.

In orthogonal and biorthogonal wavelet filters, there is some frequency overlap at the boundaries to avoid distortion, which is caused due to the abrupt end of frequency band. The small overlap ensures that the boundaries of a frequency sub-band gradually ends to 0 (cf. Fig. 1.18). The filters are also designed such that they avoid aliasing. Aliasing is a phenomena that occurs when the spectra of the two filters overlap such that the coefficients belonging to the overlapped region are indistinguishable and cannot be recovered thereby creating distortion.

The function *waveinfo('WaveletName')* in MATLAB can be used to know if a wavelet family is orthogonal or biorthogonal, or both. The Coiflet and Daubechies family used for cw-ESR and pulsed dipolar signal are both orthogonal wavelets.

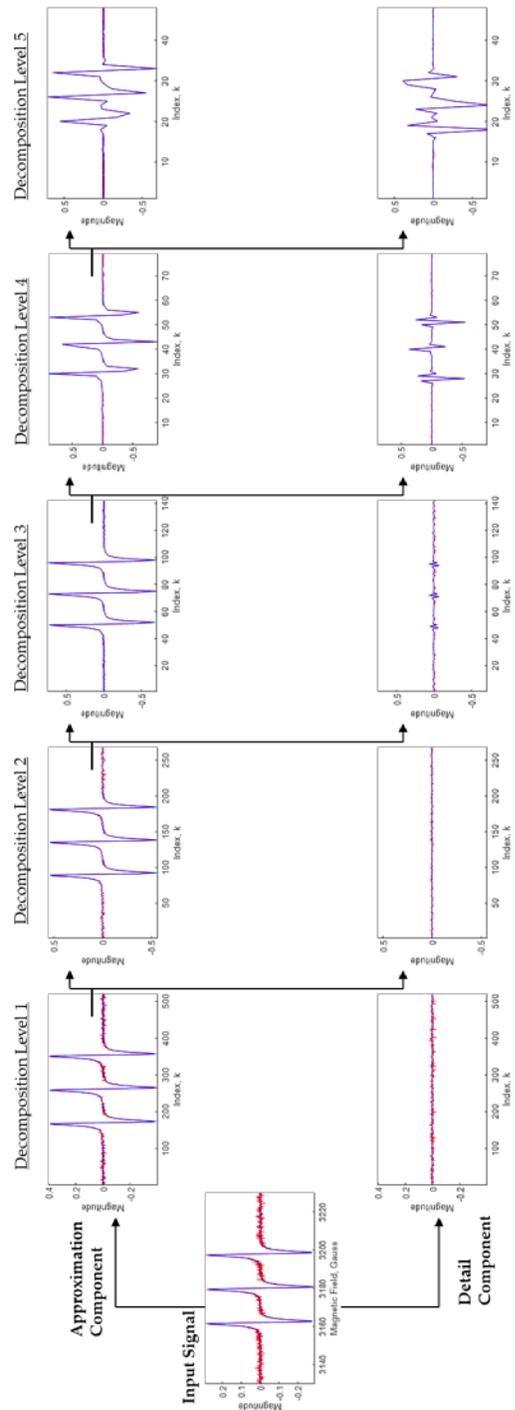


**Figure 1.18:** Comparison of non-overlapping and overlapping frequency sub-band in the wavelet domain that are used in orthogonal and biorthogonal wavelets, respectively. **a)** Non-overlapping frequency sub-band; and **b)** Overlapping frequency sub-band.

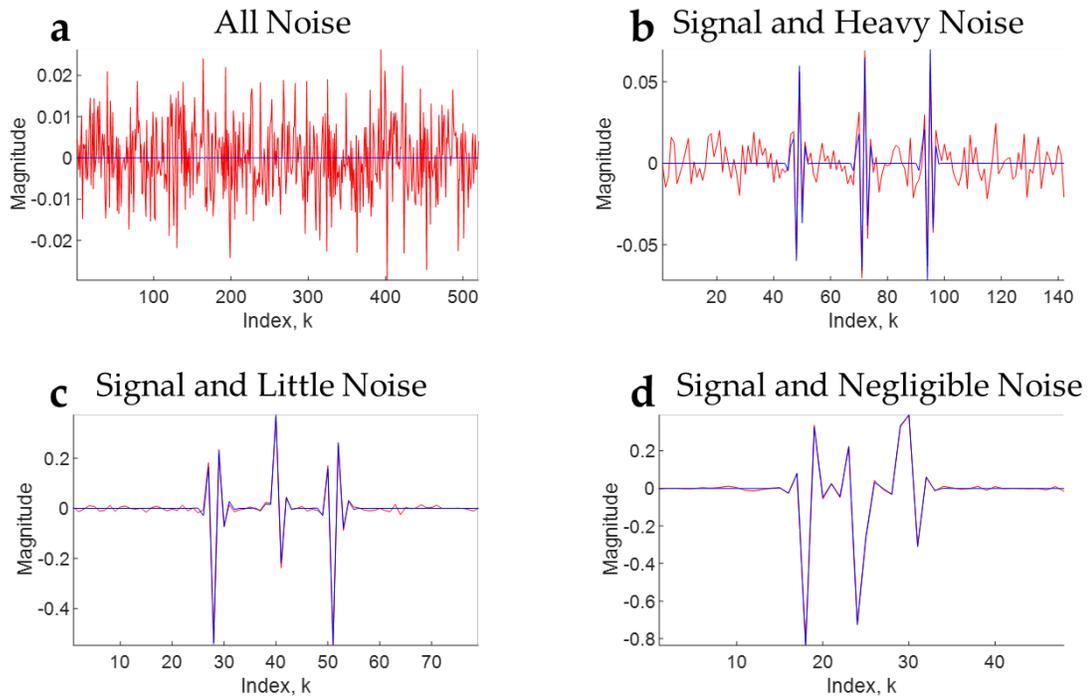
## 1.6 Wavelet Denoising – Noise Removal using Wavelet Transforms

The word “denoising” to remove noise was first used in the context of wavelet transforms<sup>5</sup>. Unlike FT and STFT, the WT can identify and distinguish noise and signal frequencies using the signal domain-frequency representation that is reflected through the Approximation and Detail components and their decomposition levels (cf. Fig. 1.19). The separation of noise from signal in WT relies on the following (cf. Fig. 1.20):

1. Noise and signal will not have the same frequency at the same signal location, because all the noise frequencies do not occur at every signal location. Therefore, when noise and signal frequencies overlap in the Approximation and Detail component, they can be distinguished using the signal location information, which is analogous to locations in the Approximation and Detail components.
2. For rare cases where noise and signal have same frequency at same signal location, the noise strength at that location is negligible compared to the signal strength and/or the contribution of the information in the overall signal is minuscule.
3. The signal strength is substantially greater than the noise strength in the wavelet domain. It is because signal is coherent and is represented by few wavelet coefficients with high magnitude. Whereas, noise being random needs to be represented by many coefficients with small magnitude. It is based on the concept of entropy: randomness has higher entropy.



**Figure 1.19:** Illustration of noisy and noise-free model signal in wavelet domain. The overlap between noisy and noise-free Detail and Approximation components is shown till decomposition level 5.



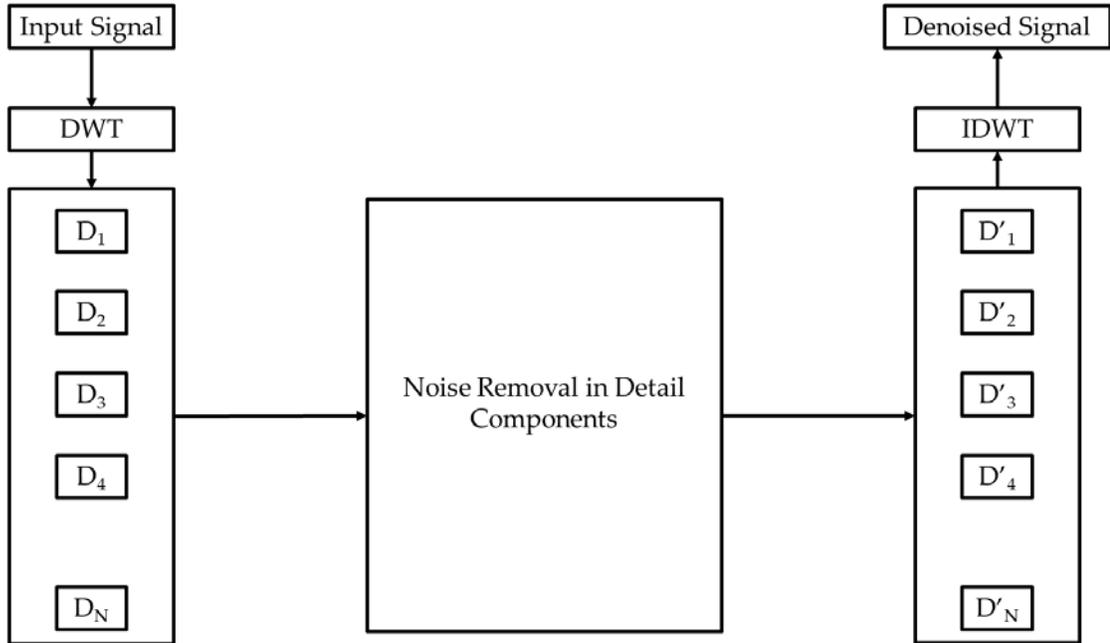
**Figure 1.20:** The four cases of the Detail component in the noisy signal. **a)** Detail component only contains noise, **b)** Detail component simultaneously contains signal and heavy noise, **c)** Detail component contains signal and little noise, and **d)** Detail component contains mostly signal with negligible noise.

### **1.6.1 Wavelet Denoising Procedure**

There are many wavelet denoising methods and approaches. However, all of them follow the following procedure:

- 1) Take a wavelet transform of the signal using a wavelet function
- 2) Process wavelet coefficients of the Detail components to remove or reduce noise.
- 3) Take the inverse wavelet transform using the same wavelet function.

Fig. 1.21 demonstrates the above procedure. Note that the Approximation component is not denoised in the standard wavelet denoising procedure.



**Figure 1.21:** The block diagram of wavelet transform based denoised procedure.  $D_j$  and  $D'_j$  are the Detail component and Denoised Detail component, respectively, at the  $j^{th}$  decomposition level.

## References

- (1) Mallat, S. *A Wavelet Tour of Signal Processing: A Sparse Way*, Third Edit.; Academic Press: New York, 2009.
- (2) Weil, J. E.; Bolton, J. R.; Wertz, J. E. *Electron Spin Resonance*; Wiley and Sons New York: Dordrecht, 1994.
- (3) Addison, P. S. *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance*; CRC Press: Philadelphia, 2002.
- (4) Daubechies, I. Ten Lectures on Wavelets. *J. Acoust. Soc. Am.* **1993**, 93 (3), 1671.
- (5) Donoho, D. L. De-Noising by Soft-Thresholding. *IEEE Trans. Inf. Theory* **1995**, 41 (3), 613–627.

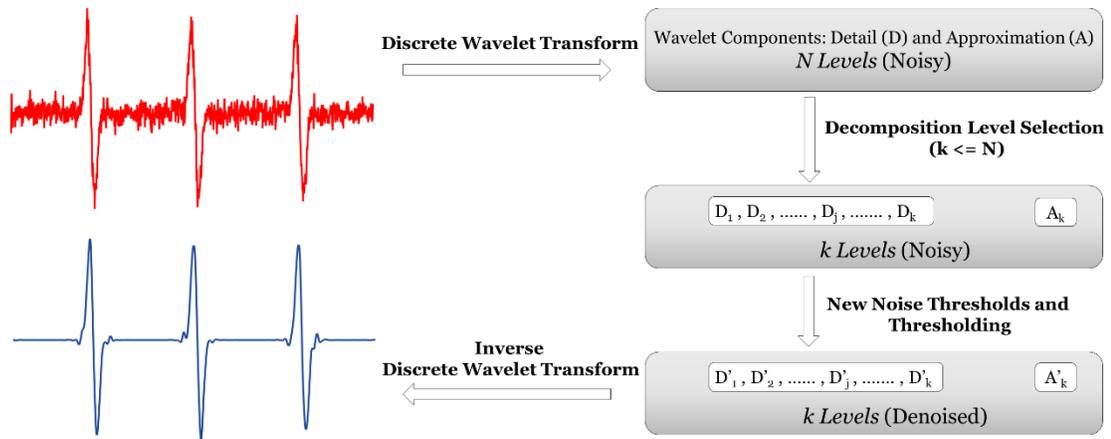
## CHAPTER 2

### A NEW WAVELET DENOISING METHOD FOR SELECTING DECOMPOSITION LEVELS AND NOISE THRESHOLDS

#### **Abstract**

A new method is presented to denoise 1-D experimental signals using wavelet transforms. Although the state-of-the-art wavelet denoising methods perform better than other denoising methods, they are not very effective for experimental signals. Unlike images and other signals, experimental signals in chemical and biophysical applications for example, are less tolerant to signal distortion and under-denoising caused by the standard wavelet denoising methods. The new method 1) provides a method to select the number of decomposition levels to denoise, 2) uses a new formula to calculate noise thresholds that does not require noise estimation, 3) uses separate noise thresholds for positive and negative wavelet coefficients, 4) applies denoising to the Approximation component, and 5) allows the flexibility to adjust the noise thresholds. The new method is applied to continuous wave electron spin resonance (cw-ESR) spectra and it is found that it increases the signal-to-noise ratio (SNR) by more than 32 dB without distorting the signal, whereas standard denoising methods improve the SNR by less than 10 dB and with some distortion. Also, its computation time is more than 6 times faster.

## Graphical Abstract



\*This chapter was published in IEEE Access journal.

**Authors:** Madhur Srivastava, C. Lindsay Anderson, and Jack H. Freed.

**Author Contributions:** Madhur Srivastava conducted research and wrote the paper, C. Lindsay Anderson contributed to the development of noise thresholds, Jack Freed supervised the research and edited the paper.

## 2.1 Introduction

Experimental signals are often difficult to study because weak signals have a low Signal-to-Noise Ratio (SNR). Based on the discrete wavelet transform (DWT), various wavelet denoising methods like wavelet shrinkage<sup>1-16</sup>, wavelet coefficient modeling<sup>17-20</sup>, and wavelet transform modulus maxima (WTMM)<sup>21-23</sup> denoising methods have been developed and shown to be more effective than filtering methods<sup>24,25</sup>.

Although denoising increases SNR, many experimentalists are skeptical of the denoised signal as they fear inadequate noise removal and/or unknown signal distortion. More importantly, the current wavelet-based denoising methods are not very reliable in accurately retrieving the signal components, especially for weak signals that have magnitude close to noise. Also, these methods try to eliminate random noise and are not tested against systematic (or coherent) noise generated by the instrument or (e.g. biological) sample. Another problem with current wavelet denoising methods is their practical implementation. They do not provide information regarding the choice of wavelets to use, the number of decomposition levels to denoise, nor do they have the flexibility to adjust the thresholds. A noisy signal clearly indicates its degree of uncertainty, whereas a denoised signal lacks such information. Therefore, to reduce noise, signal averaging<sup>26</sup> is widely used to improve experimental data. In this paper, experimental data from electron spin resonance (ESR) spectroscopy<sup>27</sup> is used, where signal averaging is currently the most reliable method to reduce noise. Although some signal denoising methods like filtering<sup>28</sup> and the

traditional wavelet denoising<sup>29</sup> methods have been applied to ESR spectra<sup>30</sup>, they have not yet yielded the desired results.

In this paper, a new wavelet denoising approach is presented which is based on wavelet shrinkage, that significantly improves denoising and provides clearer implementation compared to previous methods. It can be reliably used for denoising experimental signals. The paper is organized as follows. First, a brief description of the wavelet shrinkage method is provided. Second, the issues of estimating the choice of wavelets and accurately selecting decomposition levels to denoise are discussed. Third, the new denoising method is presented. Fourth, examples of denoising using experimental results from ESR spectroscopy are presented and compared with other wavelet shrinkage denoising methods. Finally, brief comments are given on the findings and future extensions.

## **2.2 Wavelet Shrinkage Denoising Method**

Wavelet shrinkage methods provide effective signal denoising with minimum computational complexity. In the wavelet domain, the signal is coherent and has concentrated "energy" residing in just a few high magnitude coefficients, whereas incoherent noise is represented by a large number of coefficients with small magnitudes. This sparsity of wavelet coefficients representing the signal is exploited by wavelet shrinkage methods to separate noise from signal coefficients. Figure 2.1

displays a block diagram of the denoising process. Algorithm 1 summarizes the wavelet shrinkage denoising method.

The noise threshold for the selected Detail components is obtained either using the universal threshold<sup>31</sup>  $\lambda = \sigma_j^{\widehat{Noise}} \sqrt{2 \log(N)}$  or decomposition level dependent thresholds  $\lambda_j = \sigma_j^{\widehat{Noise}} \sqrt{2 \log(N_j)}$ <sup>32</sup> or  $\lambda_j = \frac{\sigma_j^{\widehat{Noise}} \sqrt{2 \log(N_j)}}{\log(j+1)}$ <sup>33</sup>, where  $N_j$  is the length of the  $j^{th}$  Detail component, and  $\sigma_j^{\widehat{Noise}}$  is an estimate of noise level<sup>31–33</sup>. However, state-of-the-art level-dependent noise threshold selection methods like Stein's Unbiased Risk Estimate (SURE) threshold<sup>34,35</sup> and Minmax threshold<sup>31</sup> are more widely used for their better performance.

Thresholding functions are then applied to remove noise using noise thresholds. Hard and soft thresholding are the most common thresholding techniques defined in equations 2.1 and 2.2, respectively:

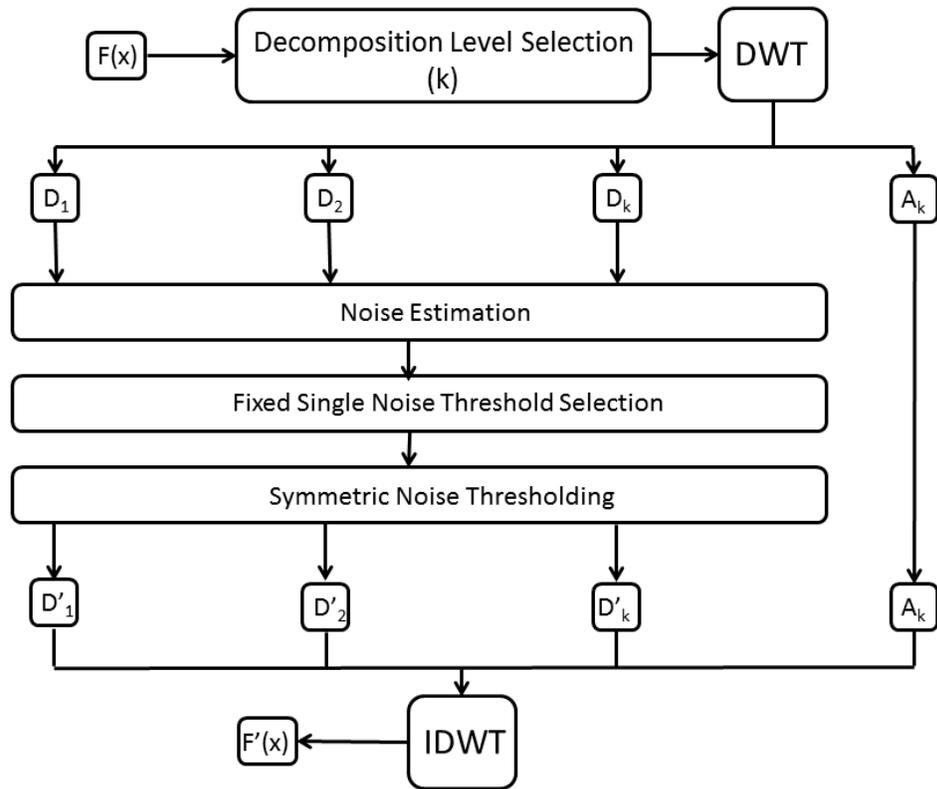
$$\widetilde{w}_{j,i} = \begin{cases} 0, & \text{for } |w_{j,i}| < \lambda_j \\ w_{j,i}, & \text{otherwise} \end{cases} \quad (2.1)$$

$$\widetilde{w}_{j,i} = \begin{cases} 0, & \text{for } |w_{j,i}| < \lambda_j \\ \text{sgn}(w_{j,i}) (|w_{j,i}| - \lambda_j), & \text{otherwise} \end{cases} \quad (2.2)$$

where  $w_{j,i}$  and  $\widetilde{w}_{j,i}$  are noisy and denoised wavelet coefficients, respectively, at the  $j^{th}$  decomposition level and the  $i^{th}$  location of the Detail component and  $j \leq k$ . Hard thresholding is better suited when a Detail wavelet coefficient is either a signal or a noise coefficient. On the other hand, soft thresholding performs better when a Detail wavelet coefficient contains both signal and noise. There are other thresholding functions in the literature that can also be used.<sup>9,36-39</sup>

### 2.2.1 Algorithm: Wavelet Shrinkage Denoising

1. Select a wavelet.
2. Select  $k$  ( $1 \leq k \leq M$ ) decomposition levels to denoise the Detail components, where  $M = \lfloor \log_2(N) \rfloor$ , and  $N = \text{length}(X)$ , and  $X$  is the discrete input signal.
3. Take the  $k^{th}$  level discrete wavelet transform (DWT) of the discrete input signal  $X$ , decomposing it into  $k$  decomposition levels, also referred to as  $k$  Detail components and the  $k^{th}$  Approximation component.
4. Estimate the noise for the  $k$  Detail components.
5. Calculate the noise threshold for the  $k$  Detail components.
6. Apply noise thresholding to the  $k$  selected Detail components.
7. Take the inverse discrete wavelet transform (IDWT) of the resultant  $k$  Detail components and the  $k^{th}$  Approximation component.



**Figure 2.1:** Block Diagram of a Standard Wavelet Shrinkage Method.

### 2.2.2 Limitations of Current Shrinkage Methods

- 1) The choice of  $k$ , the number of decomposition levels to be denoised, is arbitrary.
- 2) The choice of  $\widehat{\sigma}_j^{Noise}$  greatly influences the noise threshold  $\lambda$ , but there is no definitive way to estimate the noise value  $\widehat{\sigma}_j^{Noise}$ . Although widely used, different noise estimates yield different noise thresholds<sup>31,32,35</sup>.
- 3) Assuming white Gaussian noise (WGN), a single noise threshold is selected and applied to the magnitudes of both the negative and positive Detail coefficients. Other noise such as Poisson noise, Rician noise, and coherent noise are not considered for positive or negative bias in noise. Even for WGN, the assumption of no distribution bias is often flawed as will be discussed in section III.
- 4) The methods used result in fixed thresholds that are not adjusted. Flexibility in adjusting thresholds enables users to reach accurate or optimal thresholds for signals, especially when the wavelet coefficients of weak signals are close to the maximum magnitude of noise.

Apart from the above reasons, the current denoising methods do not emphasize the choice of wavelets that is necessary to create sparsity, and separate noise and signal in the Detail components. One possible reason is that these methods are mostly applied in image processing applications which already use customized wavelets.

## 2.3 Noise in Wavelet Domain

It is assumed that random WGN translates into random WGN in the wavelet domain. Ideally, random white gaussian noise with zero mean should be symmetric with no overall positive and negative bias, and the magnitude of minimum and maximum values should be approximately the same. Note that the minimum is the largest magnitude among negative coefficient values ( $\max(|w_{j,i} < 0|)$ ). To test this noise behavior in the wavelet domain, simulations were run in *MATLAB* using the following steps:

- 1) Generate random WGN with zero mean, call it  $X$ , and length  $N$ .
- 2) Take the full scale DWT of  $X$  that results in  $\lfloor \log_2(N) \rfloor$  Detail components.
- 3) Calculate the skewness of each Detail component. The skewness is a parameter that measures positive or negative bias in a distribution with respect to the mean. For WGN (i.e. normal distribution), the skewness should be zero.
- 4) Find the magnitude of minimum and maximum value of each Detail Component.
- 5) Calculate the percentage difference between the magnitude of minimum and maximum value with respect to the minimum value between them, i.e.

$$\frac{|\max(|w_j \geq 0|) - \max(|w_j < 0|)|}{\min(\max(|w_j \geq 0|), \max(|w_j < 0|))} \times 100$$

The percentage values should be close to zero as the magnitude of minimum and maximum values would be expected to be almost the same.

The following are the simulation parameters:

- Number of Simulations: 1000-5000.
- Noise Length  $N$ : 1024-4096.
- Wavelet Families: Coiflet, Daubechises, and Symlet.

Findings:

- The wavelet coefficient distribution for each Detail component has either positive or negative bias. The positive and negative bias is not correlated amongst successive Detail components.
- There is a substantial difference between the magnitude of minimum and maximum values, and it increases monotonically from decomposition level 1 to  $M$ . For decomposition level 1, it is already more than 10 percent.

We believe that this behavior is due to the finite nature of discrete values and the DWT. There seems to be no correlation between the choice of wavelet and the distribution bias. Experimental signals do not necessarily have simple WGN; their noise content is liable to be more complex and may have inherent positive or negative bias.

## 2.4 Wavelet Bases

It is difficult to apply the DWT without selecting an appropriate choice of wavelets. In general, standard wavelets that resemble the signal or its properties yield better signal and noise separation as well as sparsity. For example, we found that the coiflet wavelet family is better suited for cw-ESR spectra, compared to other wavelet families such as daubechies and symlets. However, a trial and error method to select the appropriate wavelet can be tedious and challenging in finding an appropriate wavelet. We recommend developing customized wavelets corresponding to the signal properties. For example, cw-ESR spectra are typically represented as blends of functions of Gaussian and Lorentzian functions<sup>27</sup>, and hence, a specific wavelet could be designed and used for such spectra.

To effectively eliminate noise from signal coefficients, the wavelets should be designed to satisfy the following conditions:

$$\max(|w_j^{Noise} \geq 0|) < \min(|w_j^{Signal} \geq 0|) \tag{2.3}$$

$$\max(|w_j^{Noise} < 0|) < \min(|w_j^{Signal} < 0|) \tag{2.4}$$

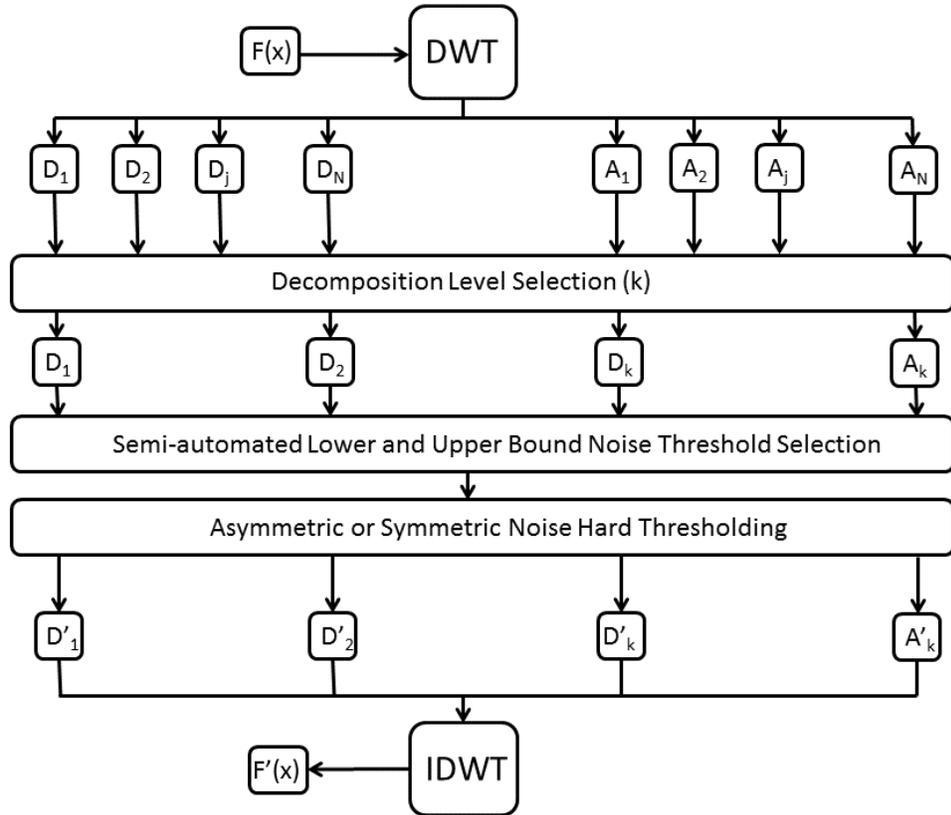
The above conditions eliminate the need for thresholding methods (like soft-thresholding) that assume a wavelet coefficient as a sum of signal and noise coefficient, i.e.  $w_{j,i} = w_{j,i}^{Signal} + w_{j,i}^{Noise}$ . In other words, either  $w_{j,i} = w_{j,i}^{Signal}$  or  $w_{j,i} = w_{j,i}^{Noise}$ .

## 2.5 New Denoising Method

The new denoising method applies the wavelet shrinkage approach with the following novel features:

- 1) The decomposition levels to denoise can be determined by the user through visual inspection of the Detail and Approximation components as described below. Alternatively, an objective measure to select the  $k$  decomposition levels is provided.
- 2) Then a new noise threshold selection formula is presented that allows the user automatic adjustment of the noise thresholds. Such noise thresholds do not require the use of a noise estimation formula.
- 3) Two distinct noise thresholds are used in the thresholding function for negative and positive wavelet coefficients, respectively.
- 4) The  $k^{th}$  Approximation component is noise thresholded along with the  $k$  Detail Components.

These features are discussed below. Figure 2.2 illustrates the new denoising method.



**Figure 2.2:** Block Diagram of the New Method

### **2.5.1 Decomposition Level Selection**

Selecting decomposition levels to denoise is a major challenge to which current methods do not provide guidance. Generally decomposition levels between 2-5 are arbitrarily selected without first examining the Detail components. A poor choice can result either in signal under-denoising or distortion, similar to that caused by poor noise thresholding. For experimental signals, this can lead to misleading analysis because the accuracy of denoising cannot be verified for every signal. A subjective and an objective method are provided to select the decomposition level  $k$ .

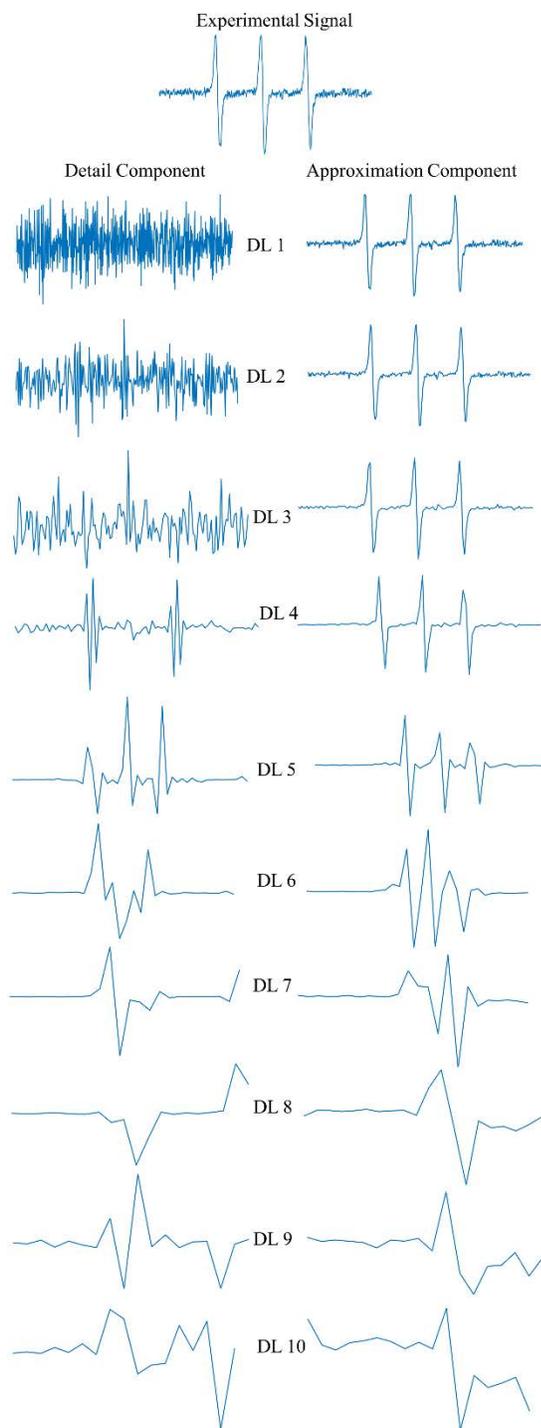
#### ***2.5.1.1 Subjective Method***

As all the Detail components and their corresponding Approximation components are correlated, visually observing all the components can easily allow one to select the correct decomposition level. In the new method, the user determines the Detail components to be denoised with the help of visual correlation. An easy way is to select decomposition levels until one is reached in which noise is almost indistinguishable. The location and magnitude of signal and noise in the wavelet component are also useful sources of information, especially for identifying systematic noise. In all the Detail components, signals occur in the same locations with large magnitudes, whereas random noise appears inconsistently with small magnitudes, and systematic noise usually occurs at a specific location with low magnitude. The amount of noise present in the Detail components reduces from decomposition level 1 to decomposition level  $M$  because noise usually contains more high frequencies than low

frequencies. For very low SNR, initial decomposition levels like 1 and 2 may contain just noise, whereas for very high SNR, the last decomposition levels may only contain signal. Through Figure 2.3, decomposition level selection is illustrated by an example from ESR. Examining the figure, one can see that the Detail components at decomposition levels 1-5 contain varying amount of noise, whereas the noise (low frequency) from decomposition level 6 is extremely small and indistinguishable. The noise at and above decomposition level 6 is represented in the Approximation component at decomposition level 5. Hence, the decomposition level  $k = 5$  can be confidently selected.

After selecting decomposition level  $k$ ,  $k$  Detail components and the  $k^{th}$  Approximation component are noise thresholded as shown in Figure 2.2. Coupled with noise thresholding, the decomposition level selection process also acts as a feedback for optimal thresholds and correct denoising by allowing visual comparison between the thresholded and non-thresholded wavelet components. Based on the feedback, the  $\kappa_{j,L}$  and  $\kappa_{j,H}$  values (discussed below) can be adjusted to desired thresholds. The standard state-of-the-art wavelet shrinkage methods do not exploit the correlation amongst wavelet components nor use this approach. They use non-adjustable thresholds which are optimized only for random Gaussian noise.

This subjective method is useful when a single set of experimental data needs to be denoised. It provides control to the experimentalist in selecting the decomposition levels as well as insights about the signal observed for analysis.



**Figure 2.3:** All the Detail and Approximation components of a cw-ESR spectrum. DL is the decomposition level. After examining the figure,  $k = 5$  can be confidently selected.

### 2.5.1.2 Objective Method

The objective measure of selecting  $k$  is more appropriate for real time denoising and where extensive experimental data needs to be denoised. To obtain the decomposition levels for noise thresholding, first calculate the "peak-to-sum ratio" ( $S_j$ ) of the Detail components,

$$S_j = \frac{\max(|w_j|)}{\sum_{i=1}^{N_j} |w_{j,i}|} \quad (2.5)$$

$S_j$  reflects the sparsity of a Detail component and allows the identification of noise presence in a Detail component. A large  $S_j$  implies signal presence with only a few large coefficient values, whereas a small  $S_j$  reveals noise presence with a large number of small coefficient values.

Now the determination of the decomposition levels  $1, 2, 3, \dots, k$  to noise threshold is described with reference to the four categories:

- 1) Detail components that only contain noise coefficients (see DL1 in Figure 2.3),  
 $S_j \leq 0.01$ .
- 2) Detail components that mainly contain noise with very few high magnitude signal coefficients (see DL2 and DL3 in Figure 2.3),  $0.01 < S_j \leq 0.1$ .
- 3) Detail components dominated by signal coefficients with noise coefficients having small magnitudes (see DL4 and DL5 in Figure 2.3),  $0.1 < S_j \leq T_r$ .

- 4) Detail components that only contain signal coefficients, i.e. noise is no longer distinguishable (see DL6 - DL10 in Figure 2.3),  $S_j > T_r$ .

Here  $T_r$  is a criterion to distinguish between the Detail components of types 3 and 4 having some noise and almost no noise. Therefore,  $k = j$  where  $S_j \leq T_r$  and  $S_{j+1} > T_r$ . We find that the choice of  $T_r \approx 0.2$  leads to an effective criterion between cases 3 and 4 (More generally,  $0 \leq T_r < 1$  given that  $S_j$  of equation 2.5 is limited by  $0 \leq S_j < 1$ ). Note that the Detail components in the above-mentioned categories occur in sequential order (see Figure 2.3). Depending on the noise level in a signal, the initial Detail component categories may or may not be present.

### **2.5.2 Threshold Selection**

Wavelet shrinkage denoising depends greatly on the noise threshold value because a poor choice of threshold can result in unremoved noise or distorted signal. The presence of signal and noise coefficients in a Detail component may not result in a symmetric distribution of coefficients and zero mean. The coefficients of signal, systematic noise, and other non-symmetric noise, like Poisson noise, can create positive or negative bias in the Detail component and result in non-zero mean. Even random Gaussian noise is likely to have coefficient distribution bias due to the Detail components' discretization and finite length as was noted in section III. In fact, it was found that symmetric Gaussian noise with zero mean has substantial differences in

magnitudes of minimum and maximum noise values in the discrete wavelet domain, and the difference increases with the decomposition level (cf. section III).

Therefore, one selects two noise thresholds ( $\lambda_L$  and  $\lambda_H$ ) in the following way,

$$\lambda_{j,L} = \mu_j - \kappa_{j,L}\sigma_j \tag{2.6}$$

$$\lambda_{j,H} = \mu_j + \kappa_{j,H}\sigma_j \tag{2.7}$$

where  $\lambda_{j,L}$  and  $\lambda_{j,H}$  are the lower and upper thresholds at decomposition level  $j$ ;  $\kappa_{j,L}$  and  $\kappa_{j,H}$  are adjustable parameters for each threshold; and  $\mu_j$  and  $\sigma_j$  are the mean and standard deviation, respectively, of the wavelet component at decomposition level  $j$ , and are defined here as,

$$\mu_j \equiv \frac{\sum_{i=1}^{N_j} w_{j,i}}{N_j} \tag{2.8}$$

$$\sigma_j \equiv \sqrt{\frac{1}{N_j - 1} \sum_{i=1}^{N_j} (w_{j,i} - \mu_j)^2} \tag{2.9}$$

The two noise thresholds in equations 2.6 and 2.7 do not require the calculation of the noise level in a Detail component. They use the mean ( $\mu_j$ ) and standard deviation ( $\sigma_j$ ) of the  $j^{th}$  Detail component, which includes both signal and noise coefficients, and  $\kappa_{j,L}$  and  $\kappa_{j,H}$  are used to accurately scale  $\sigma_j$  in eliminating noise coefficients. Also, the two thresholds include the mean value  $\mu_j$ , and do not assume it to be zero.

To obtain appropriate  $\kappa$  values, we first find the minimum values of  $\kappa_{j,L}$  and  $\kappa_{j,H}$  that cover the all the coefficients in the  $j^{th}$  Detail component. Substituting minimum (peak negative value) and maximum (peak positive) values of the Detail component in equations 2.6 and 2.7, respectively, one can express them as:

$$\kappa_{j,Lmin} = \frac{\mu_j - \max(|w_j < 0|)}{\sigma_j} \quad (2.10)$$

$$\kappa_{j,Hmin} = \frac{\max(|w_j > 0|) - \mu_j}{\sigma_j} \quad (2.11)$$

Since all the decomposition levels  $j$  ( $1 \leq j \leq k$ ) where  $S_j \leq 0.01$  (i.e., case 1) contain only noise coefficients, for these cases we can set  $\kappa_{j,L} = \kappa_{j,Lmin}$  and  $\kappa_{j,H} = \kappa_{j,Hmin}$ . In other words, all the coefficients of the Detail component are assigned zero.

For all the decomposition levels  $j$  ( $1 \leq j \leq k$ ) where  $0.01 < S_j < T_r$  (i.e., cases 2 and 3), both signal and noise coefficients are present. Knowing that the magnitude of a noise coefficient is less than that of the signal coefficients (cf. equations 2.3 and 2.4),  $\kappa_{j,L} < \kappa_{j,L_{min}}$  and  $\kappa_{j,H} < \kappa_{j,H_{min}}$ . One then obtains  $\kappa_{j,L}$  and  $\kappa_{j,H}$  from  $\kappa_{j,L_{min}}$  and  $\kappa_{j,H_{min}}$  using the peak positive and negative coefficient values in the following way,

$$\kappa_{j,L} = \left( \frac{S_{r,L} - S_{j,L}}{S_{r,L}} \right) \kappa_{j,L_{min}} \quad (2.12)$$

$$\kappa_{j,H} = \left( \frac{S_{r,H} - S_{j,H}}{S_{r,H}} \right) \kappa_{j,H_{min}} \quad (2.13)$$

where  $S_{r,L}$  and  $S_{r,H}$  are the reference peak-to-sum coefficient values and are defined as  $S_{r,L} \equiv \frac{S_{k,L} + S_{k+1,L}}{2}$  and  $S_{r,H} \equiv \frac{S_{k,H} + S_{k+1,H}}{2}$ .  $S_{j,L}$  and  $S_{j,H}$  are the peak-to-sum ratios of the negative and positive coefficient values, respectively. They are calculated from equation 2.5 as,

$$S_{j,L} = \frac{\max(|w_j < 0|)}{\sum_{i=1}^{N_j} |w_{j,i} < 0|} \quad (2.14)$$

$$S_{j,H} = \frac{\max(|w_j \geq 0|)}{\sum_{i=1}^{N_j} |w_{j,i} \geq 0|} \quad (2.15)$$

By separating the wavelet coefficients at a decomposition level into two groups, one having negative values and other having positive values, the number of coefficient values as well as their sum in each category is reduced. However, the peak value remains at a similar level. In other words, one expects  $\max(|w_j < 0|) \approx \max(|w_j \geq 0|)$ , and one of them is in fact  $\max(|w_j|)$ , whereas  $\sum_{i=1}^{N_j} |w_{j,i} < 0|$  and  $\sum_{i=1}^{N_j} |w_{j,i} \geq 0|$  are both less than their sum  $\sum_{i=1}^{N_j} |w_{j,i}|$ . Thus  $S_{j,L}; S_{j,H} > S_j$ . Instead of manually selecting the  $\kappa$  values, it is found that equations 2.12 and 2.13 accurately estimate the required  $\kappa_{j,L}$  and  $\kappa_{j,H}$  values.

The Detail components at decomposition levels  $j > k$  (i.e.,  $S_j > T_r$ ) are not noise thresholded as they contain only signal coefficients or signal and indistinguishable low frequency noise coefficients.

Although  $\kappa_{j,L}$  and  $\kappa_{j,H}$  will optimally select the noise thresholds, they may require minor adjustment by user intervention for the following three exceptions:

- 1) Even for  $S_j \geq 0.01$ , a Detail component can have only noise coefficients due to one or more large random noise spikes. In this case,  $\kappa_{j,L} = \kappa_{j,L_{min}}$  and  $\kappa_{j,H} = \kappa_{j,H_{min}}$ , instead of equation 2.12 and 2.13.

- 2) In a noise-dominated Detail component, either positive or negative coefficients are all noise, requiring readjustment  $\kappa_{j,L} = \kappa_{j,L_{min}}$  or  $\kappa_{j,H} = \kappa_{j,H_{min}}$ , respectively.
- 3) When small signal and large noise coefficient values are comparable,  $\kappa_{j,L}$  and/or  $\kappa_{j,H}$  may require a slight increase or decrease in their values to separate signal and noise coefficients.

### 2.5.3 Thresholding Function

This new approach to denoise the signal uses hard thresholding coupled with the above noise threshold selection method. Unlike other thresholding techniques, hard thresholding ensures that the signal wavelet coefficients are not distorted by the noise thresholds, especially when signal and noise coefficients are disjoint and not superimposed. With appropriate wavelet selection, the noise present at a signal location is separated in the initial decomposition levels and is removed using noise thresholding. Hard thresholding is now carried out in the following way,

$$\widetilde{w}_{j,i} = \begin{cases} 0 & : \lambda_{j,L} \leq w_{j,i} \leq \lambda_{j,H} \\ w_{j,i}, & : otherwise \end{cases} \quad (2.16)$$

$w_{j,i}$  is the  $i^{th}$  Detail coefficient at  $j^{th}$  decomposition level,  $\widetilde{w}_{j,i}$  is the thresholded wavelet coefficient, and  $\lambda_{j,L}$  and  $\lambda_{j,H}$  are the two noise thresholds obtained from equations 2.6 and 2.7. The above equation is a generalized form of the hard

thresholding function shown in equation 2.1. If  $\kappa_{j,L} = \kappa_{j,H}$  and  $\mu = 0$ , then  $\lambda_{j,L} = \lambda_{j,H} = \lambda_j$ , and equation 2.16 becomes equation 2.1.

#### **2.5.4 Denoising the Approximation Component**

Current wavelet shrinkage methods do not denoise the Approximation component because the noise thresholding relies on sparsity of wavelet coefficients for estimating noise and selecting thresholds. Experimental signals often contain substantial low frequency noise, especially at low SNR, that impedes analysis. The Detail components that represent low frequency noise are non-sparse, restricting the effectiveness of noise thresholding. However, the Approximation component can also contain low frequency noise. Thus, it would be desirable to remove the Approximation component's low frequency noise. Depending on the signal properties and the choice of wavelet, the  $k^{th}$  Approximation component can represent low frequency noise and signal coefficients in a relatively sparse manner. The signal coefficients will have larger values than the noise coefficients. This new approach denoises the  $k^{th}$  Approximation component to eliminate low frequency noise. Similar to the Detail components, the noise thresholds for the Approximation component is selected using the Approximation coefficients in equations 2.6 and 2.7, where  $\kappa$  values are obtained from equations 2.12 and 2.13. The thresholding function in equation 2.16 is then applied to remove the noise. This is shown in the block diagram of Figure 2.2.

## 2.6 ESR Experiments and Results

### 2.6.1 ESR Methodology and Sample Preparation

Most of the ESR experiments were performed at 20°C on a commercial spectrometer (*BRUKER ELEXYS-II E500*) at a standard microwave frequency of 9.4 GHz (X-band) corresponding to a dc magnetic field of 0.34 Tesla (cf. Figs. 2.3-2.8). These results are referred to as example 1. In these experiments the microwave frequency is held constant, while the magnetic field is swept through resonance to obtain the ESR spectrum<sup>27</sup>. The sample consisted of 4  $\mu\text{L}$  of a 100  $\mu\text{M}$  aqueous solution of the commonly used spin-probe molecule Tempol (*4-Hydroxy-2,2,6,6-Tetramethylpiperidine 1-oxyl*).<sup>27,40-43</sup> It was placed in a glass capillary of 0.8 mm ID, which was then introduced into the microwave cavity situated between the pole caps of the dc magnet. The magnetic field was then swept over a range of 60 G corresponding to the resonant spectral range which took 2 minutes, and a 82 ms time constant was used. The spectral data consisted of 4096 points along the magnetic-field sweep. In addition, small coils placed at the sides of the resonator provided a small magnetic field modulation of  $\pm 0.02$  G at a frequency of 100 kHz. The 100 kHz modulated ESR signal was detected with a lock-in detector at this frequency, providing the first derivative of the absorption signal.<sup>27,44</sup> Low power (0.2 mW) microwave radiation was used to avoid saturating the ESR signal. Multi-scan experiments were performed with a delay of 4 s between scans. The results of these scans were then averaged. The ESR spectrum in Figure 2.3 was obtained under very similar conditions, also for the Tempol spin-probe.

In example 2, an ESR spectrum (Figure 2.9) was obtained under different conditions to provide a more complex spectrum for denoising. It was obtained on a home-built (ACERT) 95 GHz ESR spectrometer<sup>42</sup> with a dc magnetic field of 3.3 Tesla at 25°C. The sample here contained ca. 5  $\mu$ L of phospholipid vesicles doped with 0.5% of a lipid spin label: 16-PC (1-acyl-2-[16-(4,4-dimethyloxazolidine-N-oxyl)stearoyl]-sn-glycero-3-phosphocholine) in the fluid phase that had been suspended in water. It was placed in a disc-like sample holder utilized for millimeter-wave ESR methodology.<sup>42,45</sup> The acquisition parameters were: sweep width of 250 G, sweep time of 2 min with a time constant of 100 ms. The millimeter-wave power was 16 mW and the spectrum consists of 512 points. The field modulation parameters were: 6 G modulation amplitude and 100 kHz modulation frequency. The spectrum in Figure 2.9 is the average of 18 scans. The time between scans was 3 s.

### **2.6.2 Experiment**

These experimental cw-ESR spectra were used to test and compare the denoising methods. By averaging different numbers of scans, different SNRs were obtained. In the case of typical cw-ESR spectra, such as those in Figures 2.3-2.9, they have regions with zero signal, aiding in the estimation of noise. The noise in these experiments is predominantly random Gaussian noise, as is typical in ESR. Hence, the cw-ESR spectra reveal the effectiveness of denoising methods at places where the signal is present as well as absent. In Figure 2.4 two noisy spectra from a single sample are

shown, but at different SNRs obtained by signal averaging 4 and 16 scans, respectively. As a reference, the signal averaged over 500 scans is used for comparison with the denoised signals. The SNR was calculated as,

$$SNR = \frac{Signal_{Peak-to-Peak}}{Noise_{rms}} \quad (2.17)$$

or in decibels (dB)

$$SNR = 20 \log_{10} \left( \frac{Signal_{Peak-to-Peak}}{Noise_{rms}} \right) \quad (2.18)$$

where *rms* is the root mean square. The SNR measures both distorting (i.e., structural) and non-distorting (i.e., non-structural) noise in the signal, but it cannot differentiate between them. Usually, the structural fidelity of the noisy or denoised signal with respect to a reference signal can be visually observed. Of course, the retention of signal structure is important. Apart from visually observing and comparing the noisy or denoised signal with a reference signal, an objective measure, the structure similarity index measure (SSIM)<sup>46</sup> was used. It enabled the estimation of the structural similarity or fidelity of the noisy and of the denoised signals at 4 and 16 scans with respect to the reference 500 scan signal. The SSIM is calculated as,

$$SSIM(X, Y) = \frac{(2\mu_X\mu_Y + c_1)(2\sigma_{XY} + c_2)}{(\mu_X^2 + \mu_Y^2 + c_1)(\sigma_X^2 + \sigma_Y^2 + c_2)} \quad (2.19)$$

where  $X$  is either the noisy or denoised signal;  $Y$  is the reference signal;  $\mu_X$  and  $\mu_Y$  are the mean values of  $X$  and  $Y$ , respectively;  $\sigma_X$  and  $\sigma_Y$  are the standard deviation values of  $X$  and  $Y$ , respectively;  $\sigma_{XY}$  is the covariance of  $X$  and  $Y$ ; and  $c_1$  and  $c_2$  are small positive constants used for stabilizing each term. The SSIM value ranges within  $[-1,1]$ , and achieves 1 when  $X$  is identical to  $Y$  ( $SSIM(X,Y) = SSIM(X,X) = SSIM(Y,Y) = 1$ ). The more that  $X$  structurally resembles  $Y$ , the SSIM value will be closer to 1. *MATLAB* was used to calculate the SSIM.

An ESR spectrum is frequently composed of Lorentzian and Gaussian functions, or mixtures of both. The examples of Figures 2.3-2.8 are simple Lorentzians, whereas that of Figure 2.9 is more complex. If the cw-ESR spectra match the wavelets, the denoising can be extremely effective. Thus, we used *coiflet 3* as the wavelet because it best resembles the spectra. Other *coiflets* and wavelets such as *daubechies* and *symlet* did not perform as well as *coiflet 3*.

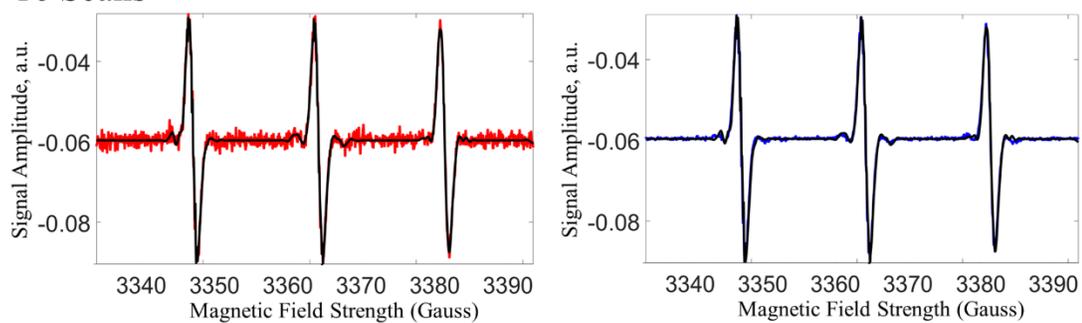
The new method is compared with the standard *SUREShrink* and *Minimax* methods<sup>29,31</sup> using both hard and soft thresholding as well as other wavelet denoising methods. We used the two different examples (Figures 2.5-2.8 and Figure 2.9) to show and compare the denoised results. The *SUREShrink* and *Minimax* methods are the optimized hybrid scale dependent threshold selection schemes in the current literature<sup>24,47</sup>, whereas other newer methods<sup>9,37-39</sup> use a thresholding function other than hard and soft in an effort to obtain better denoising. Filtering methods are not compared with the new method because standard wavelet denoising methods perform

better than them as shown in ESR<sup>30</sup> and in general<sup>24,25</sup>. The decomposition level  $k$  that yields maximum SNR compared to other decomposition levels is selected for the standard denoising methods. In the new method,  $T_r = 0.20$  was used for all the examples. We have extensively tested (both experimental and model data sets) this value of  $T_r$  as the criterion for noise thresholding and find it succeeds in virtually all cases.

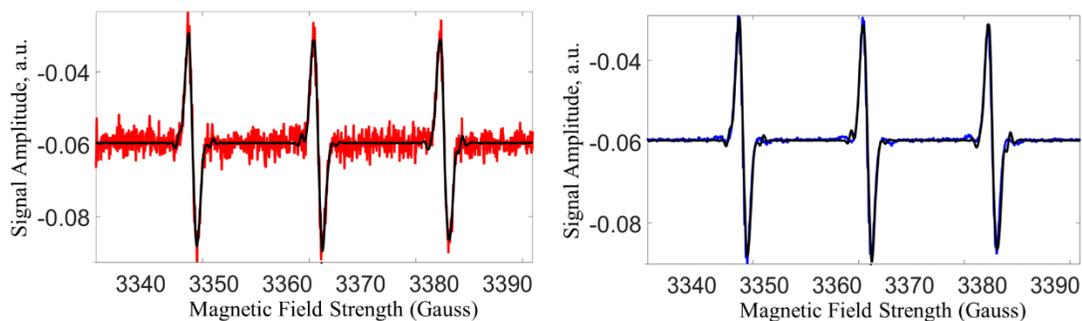
### **2.6.3 Results**

In example 1 (cf. Figures 2.3-2.8), the denoised spectra by the new method at 16 and 4 scans in Figure 2.4 are shown overlapped with the noisy and reference spectra for comparison and analysis. As can be seen, the new method successfully recovers the signal peaks and removes the baseline noise. Also, the new method retrieves the small satellite peaks (adjacent to large peaks) that were submerged in noise. Comparing the spectra of 16 and 4 scans, the denoised 16 scan spectrum recovers the satellite peaks more accurately than the 4 scan denoised spectrum, as expected. In Figures 2.5-2.8 the new method is compared with the denoised spectra of other denoising methods.

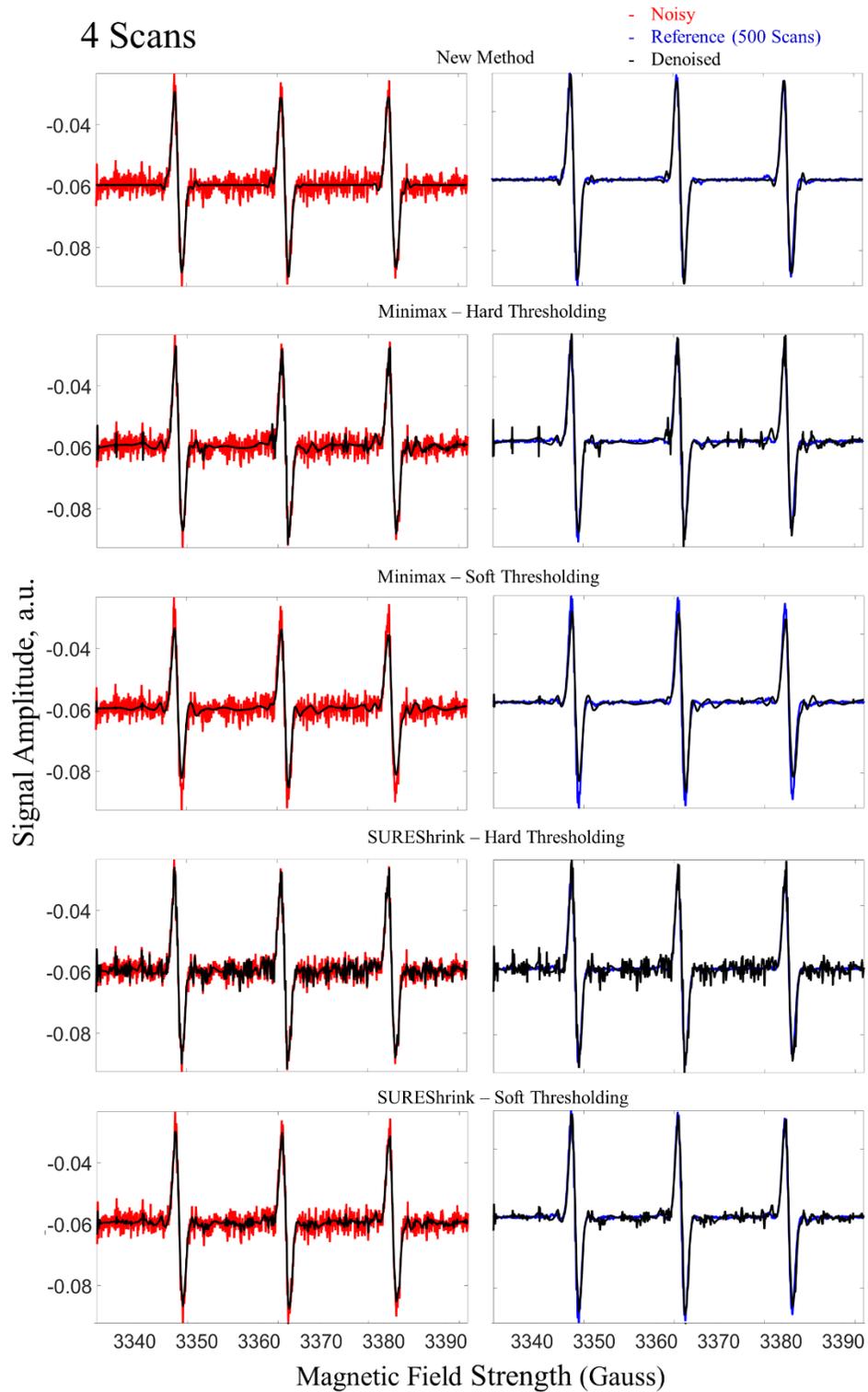
### 16 Scans



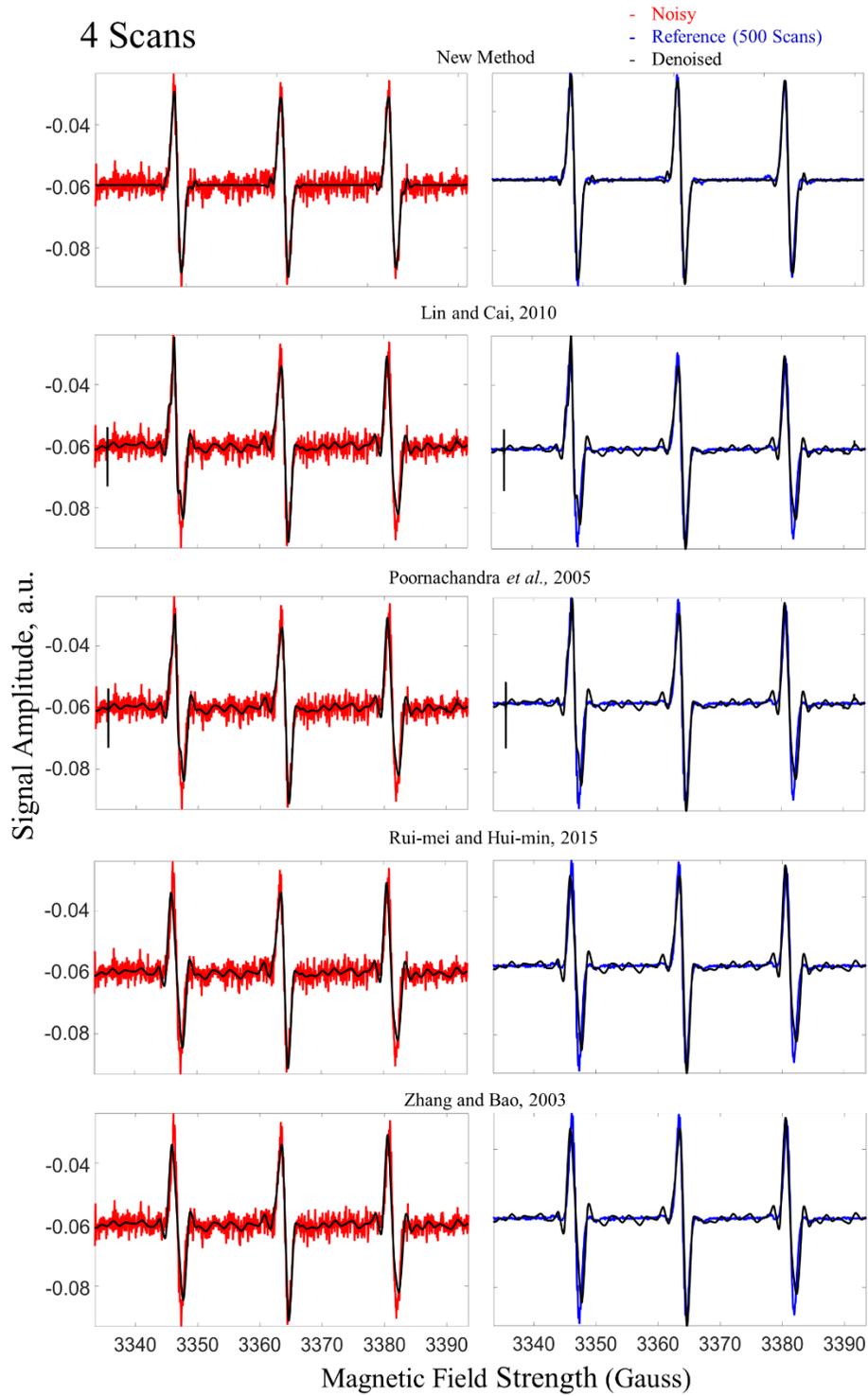
### 4 Scans



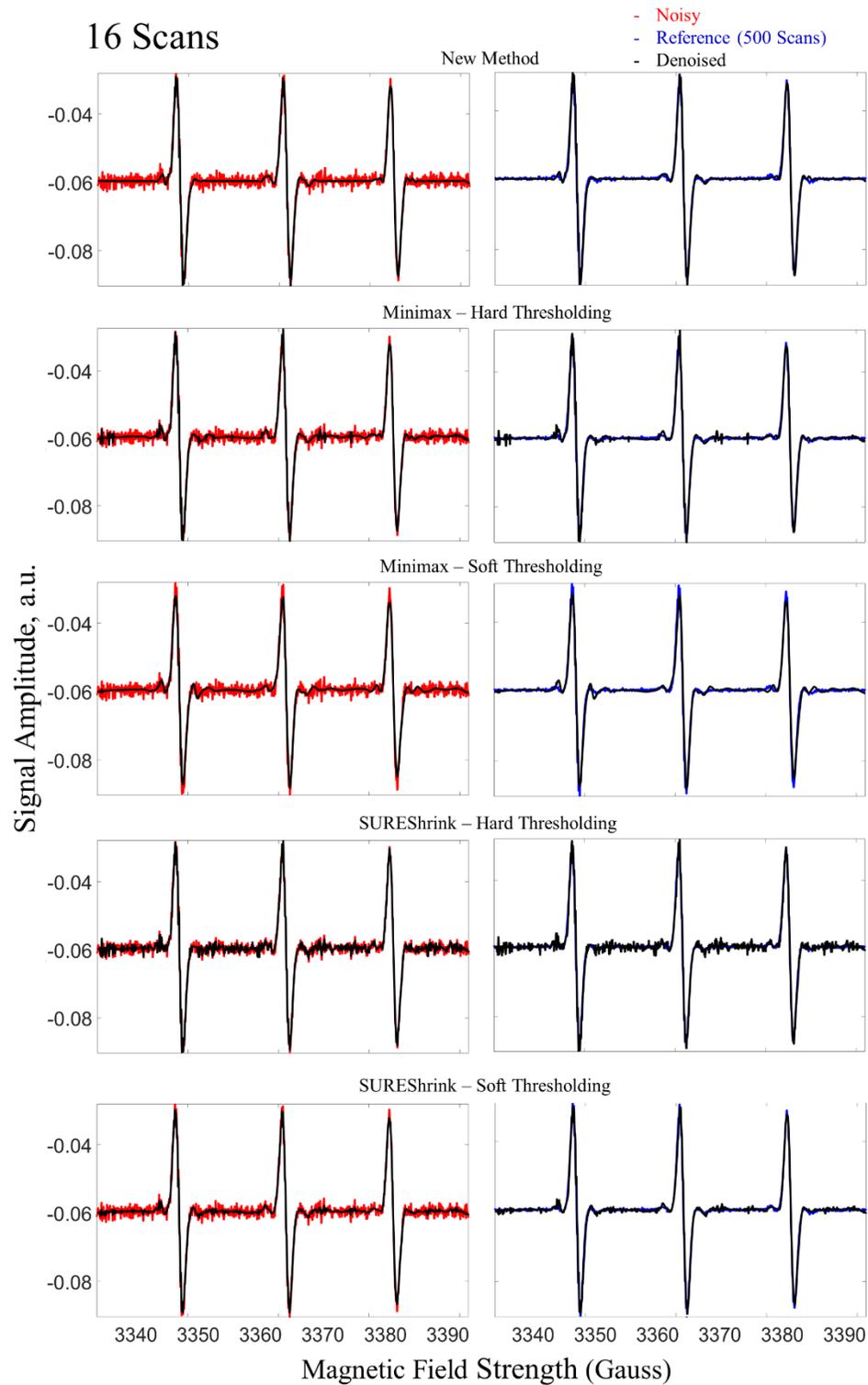
**Figure 2.4:** Example 1 – New Method Denoising results at 16 and 4 scans, respectively. The cw-ESR spectrum was obtained from a sample of the nitroxide spin probe tempol in a 0.1 mM solution in water.<sup>48</sup> The signal amplitude is in arbitrary units (a.u.) and the abscissa is in Gauss.



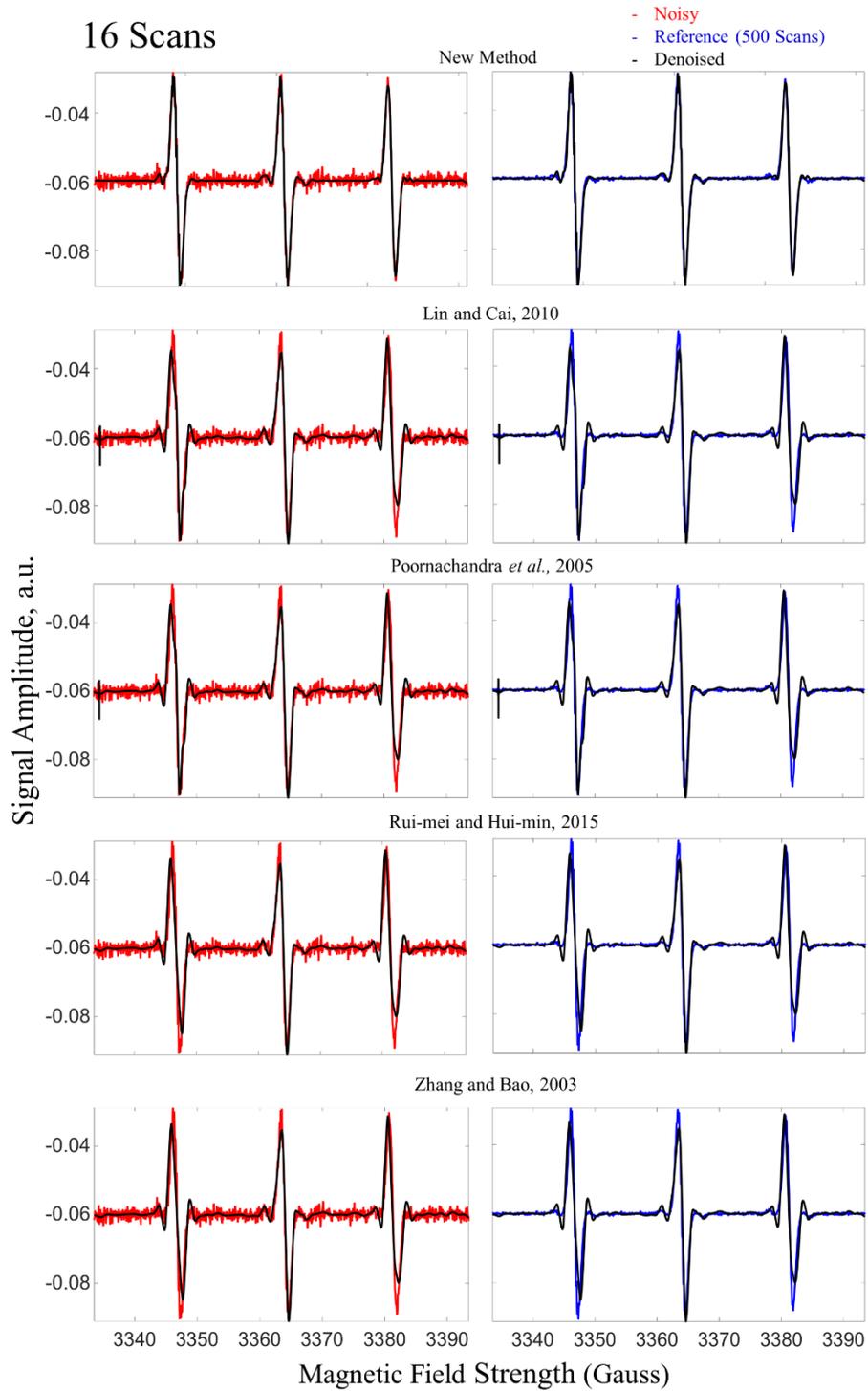
**Figure 2.5:** Experimental Data Example 1 – Comparison of New Method with Minimax and SUREShrink Denoising at 4 scans.



**Figure 2.6:** Example 1 – Comparison of New Method with Denoising methods<sup>9,37–39</sup> other than soft and hard thresholding at 4 scans.

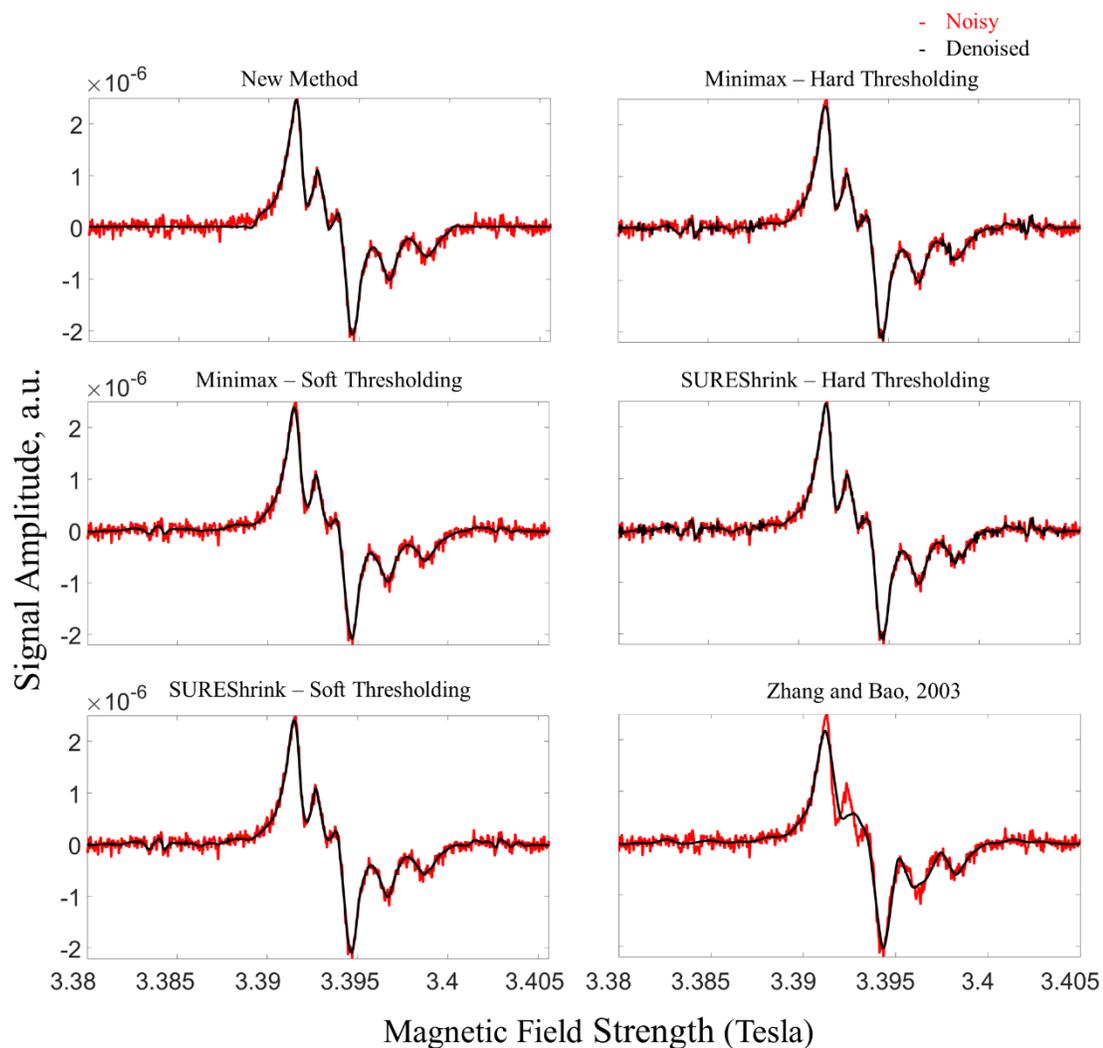


**Figure 2.7:** Example 1 – Comparison of New Method with Minimax and SUREShrink Denoising at 16 scans.



**Figure 2.8:** Example 1 – Comparison of New Method with Denoising methods<sup>9,37–39</sup> other than soft and hard thresholding at 16 scans.

Similarly in example 2 (see Figure 2.9), the new method effectively removes the baseline noise and is better able to retrieve the peaks compared to other denoising methods. Note that the results of non-soft and non-hard thresholding function methods<sup>9,37-39</sup> were similar for example 2, and hence, only Zhang and Song's method<sup>39</sup> is shown. As can be seen, the current state-of-the-art methods are either not very effective at removing baseline noise and/or they distort the signal peak heights.



**Figure 2.9:** Example 2 – Comparison of New Method with other standard denoising methods. The cw-ESR spectrum was obtained from a sample of a nitroxide-labeled lipid in lipid vesicles.

Tables 2.1 and 2.2 show the SNR of the original noisy and denoised spectra, along with decomposition levels selected for noise thresholding. In example 1, it can be seen that the new method doubles the SNR (in dB) from the noisy spectrum and is more than 20 dB greater than other denoising methods. In example 2, which is a more complex but much broader spectrum, the new method increases the SNR by a huge amount, almost reaching the noiseless state. This is likely due to the low frequency nature of the spectrum versus the higher frequency noise.

For example 1, since there is a good reference signal, the SSIM is also calculated and shown in Table 2.1. It can be seen that the noisy signal at 16 scans is structurally very similar to the reference signal, revealing that the noise hardly distorts the signal. But the 4 scan noise possesses some structural distortion. For 16 scans, the new method along with SUREShrink and Minimax soft and hard thresholding methods increases the SSIM, while the other methods decrease it. The increase is highest for the new method. For 4 scans, the new method significantly increases the SSIM, while other denoising methods slightly improve it. At both 16 and 4 scans, the new method not only regains the reference signal maximally but it also resembles it very closely.

Both qualitatively and quantitatively, the new method outperforms the current denoising methods and successfully recovers the desired spectrum.

**Table 2.1:** Example 1 – Signal to Noise Ratio (SNR): Comparison of New Method with other standard wavelet denoising methods at 16 and 4 scans. Hard and Soft represent hard and soft thresholding, respectively; dB is SNR in decibels; and DL is Decomposition Level (cf. Figures 2.3-2.8).

Method	16 scans				4 scans			
	SNR (dB)	SNR	SSIM	DL	SNR (dB)	SNR	SSIM	DL
Noisy	35	57	0.9903	-	30	32	0.9684	-
Minimax-Hard	43	146	0.9959	7	38	81	0.9880	7
Minimax-Soft	46	201	0.9954	7	41	114	0.9890	12
SUREShrink-Hard	28	82	0.9938	9	32	41	0.9779	9
SUREShrink-Soft	44	157	0.9966	7	37	77	0.9905	12
Zhao [9]	46	209	0.9821	6	38	80	0.9790	6
Poornachandra [38]	44	167	0.9825	6	37	73	0.9803	6
Zhang [39]	46	209	0.9821	6	38	80	0.9778	6
Lin [37]	44	167	0.9825	6	38	79	0.9800	6
<b>New Method</b>	<b>73</b>	<b>4438</b>	<b>0.9969</b>	<b>6</b>	<b>64</b>	<b>1567</b>	<b>0.9957</b>	<b>6</b>

**Table 2.2:** Example 2 – Signal to Noise Ratio (SNR): Comparison of New Method with other standard wavelet denoising methods. dB is SNR in decibels; and DL is Decomposition Level (cf. Figure 2.9).

<b>Method</b>	<b>SNR (dB)</b>	<b>SNR</b>	<b>DL</b>
Noisy	33	49	-
Minimax-Hard	38	81	10
Minimax-Soft	43	149	6
SUREShrink-Hard	37	76	7
SUREShrink-Soft	41	115	9
Zhang [39]	44	167	6
<b>New Method</b>	<b>165</b>	<b><math>191 \times 10^6</math></b>	<b>4</b>

#### 2.6.4 Computational Complexity

For the new method as well as the other methods, the computational complexity of DWT, IDWT, plus application of the thresholding function is  $\mathcal{O}(N)$ . However, the new method takes  $\mathcal{O}(N)$  operations to calculate thresholds, as it requires only simple arithmetic operations for calculating mean, standard deviation, and ratios. On the other hand, the thresholds obtained from SURE and Minimax optimization can take  $\mathcal{O}(N \log N)$  operations. Therefore, overall the new method has computational complexity ( $\mathcal{O}(N)$ ) which is less than the other denoising methods ( $\mathcal{O}(N \log N)$ ).

Table 2.3 shows the actual computation times of the new method and other denoising methods for example 1. The computation time was calculated in *MATLAB* on a 64-bit operating system with 16 GB RAM and a 3.30 GHz processor. To overcome any processing variation in the computer, the computation time was measured 5000 times and then averaged. It can be seen that the new method is more than 6 times faster than other denoising methods.

**Table 2.3:** Example 1 - Computation Time: Comparison of New Method with other standard wavelet denoising methods.

Method	Computation Time (seconds)	
	16 scans	4 scans
Noisy	35	57
Minimax-Hard	43	146
Minimax-Soft	46	201
SUREShrink-Hard	28	82
SUREShrink-Soft	44	157
Zhao [9]	46	209
Poornachandra [38]	44	167
Zhang [39]	46	209
Lin [37]	44	167
<b>New Method</b>	<b>73</b>	<b>4438</b>

## 2.7 Conclusions

This paper presents a comprehensive framework for wavelet transform denoising illustrating it with denoising of cw-ESR spectra, and it shows the limitations of current wavelet denoising methods. The two examples were chosen as illustrative from over 100 examples of model and experimental data. This new denoising method enables the selection of appropriate thresholds for noise removal. It is shown that this new method is able to effectively denoise and especially retrieve the signal peaks accurately.

There are two advantages of the proposed method. First, it allows the analysis of a low SNR signal by effectively retrieving the relevant information. This is especially beneficial where signals are obtained at low SNR. Second, the new method can be used to minimize the number of scans needed for signal averaging, resulting in large savings in experimental time. In the application of the new method to extensive cw-ESR spectra, we found it to be consistently successful in accurately retrieving the original signal. Theoretically, very extensive signal averaging would remove all the random noise, but due to time constraints, the number of scans that are practical is often limited.

We believe that our new denoising method can be applied to other spectroscopies such as nuclear magnetic resonance (NMR) spectroscopy<sup>49</sup>, Infrared spectroscopy<sup>49</sup>, Ultraviolet spectroscopy<sup>49</sup>, Raman spectroscopy<sup>50</sup>, mass spectroscopy<sup>49</sup>, and others that yield 1D spectra. We currently are developing this approach for time domain signals as well as for two- and higher-dimensional spectroscopies.

## References

- (1) Cai, T.; Silverman, B. Incorporating Information on Neighbouring Coefficients into Wavelet Estimation. *Sankhyā Indian J. Stat. Ser. B* **2001**, *63* (2), 127–148.
- (2) Jun Jiang; Jian Guo; Weihua Fan; Qingwei Chen. An Improved Adaptive Wavelet Denoising Method Based on Neighboring Coefficients. In *2010 8th World Congress on Intelligent Control and Automation*; IEEE, 2010; pp 2894–2898.
- (3) Chen, G. Y.; Bui, T. D. Multiwavelets Denoising Using Neighboring Coefficients. *IEEE Signal Process. Lett.* **2003**, *10* (7), 211–214.
- (4) Chen, G. Y.; Bui, T. D.; Krzyzak, A. Image Denoising Using Neighbouring Wavelet Coefficients. *Integr. Comput. Aided. Eng.* **2005**, *12* (1), 99–107.
- (5) Dengwen, Z.; Wengang, C. Image Denoising with an Optimal Threshold and Neighbouring Window. *Pattern Recognit. Lett.* **2008**, *29* (11), 1694–1697.
- (6) Madhu, S.; Bhavani, H. B.; Sumathi, S.; Vidya, H. A. A Novel Algorithm for Denoising of Simulated Partial Discharge Signals Using Adaptive Wavelet Thresholding Methods. In *2015 2nd International Conference on Electronics and Communication Systems (ICECS)*; IEEE, 2015; pp 1596–1602.
- (7) Hussein, R.; Shaban, K. B.; El-Hag, A. H. Histogram-Based Thresholding in Discrete Wavelet Transform for Partial Discharge Signal Denoising. In *2015 International Conference on Communications, Signal Processing, and their Applications (ICCSPA'15)*; IEEE, 2015; pp 1–5.
- (8) Cetin, A. E.; Tofighi, M. Projection-Based Wavelet Denoising [Lecture Notes]. *IEEE Signal Process. Mag.* **2015**, *32* (5), 120–124.
- (9) Zhao, R.-M.; Cui, H. Improved Threshold Denoising Method Based on Wavelet Transform. In *2015 7th International Conference on Modelling, Identification and Control (ICMIC)*; IEEE, 2015; pp 1–4.
- (10) Ding, Y.; Selesnick, I. W. Artifact-Free Wavelet Denoising: Non-Convex Sparse Regularization, Convex Optimization. *IEEE Signal Process. Lett.* **2015**, *22* (9), 1364–1368.

- (11) Qiming Zhao; Wenzhan Dai. A Wavelet Denoising Method of New Adjustable Threshold. In *2015 IEEE 16th International Conference on Communication Technology (ICCT)*; IEEE, 2015; pp 684–688.
- (12) Madadi, Z.; Anand, G. V.; Premkumar, A. B. Signal Detection in Generalized Gaussian Noise by Nonlinear Wavelet Denoising. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2013**, *60* (11), 2973–2986.
- (13) Deng, G.; Liu, Z. A Wavelet Image Denoising Based on the New Threshold Function. In *2015 11th International Conference on Computational Intelligence and Security (CIS)*; IEEE, 2015; pp 158–161.
- (14) Bibina, V. C.; Viswasom, S. Adaptive Wavelet Thresholding & Joint Bilateral Filtering for Image Denoising. In *2012 Annual IEEE India Conference (INDICON)*; IEEE, 2012; pp 1100–1104.
- (15) Guo, S.; Wu, F.; Wei, W.; Guo, J.; Ji, Y.; Wang, Y. A Novel Wavelet Denoising Method Used for Droplet Volume Detection in the Microfluidic System. In *2013 IEEE International Conference on Mechatronics and Automation*; IEEE, 2013; pp 1732–1737.
- (16) Chen, G.; Xie, W.; Zhao, Y. Wavelet-Based Denoising: A Brief Review. In *2013 Fourth International Conference on Intelligent Control and Information Processing (ICICIP)*; IEEE, 2013; pp 570–574.
- (17) Chang, S. G.; Bin Yu; Vetterli, M. Spatially Adaptive Wavelet Thresholding with Context Modeling for Image Denoising. *IEEE Trans. Image Process.* **2000**, *9* (9), 1522–1531.
- (18) Sendur, L.; Selesnick, I. W. Bivariate Shrinkage Functions for Wavelet-Based Denoising Exploiting Interscale Dependency. *IEEE Trans. Signal Process.* **2002**, *50* (11), 2744–2756.
- (19) Pizurica, A.; Philips, W. Estimating the Probability of the Presence of a Signal of Interest in Multiresolution Single- and Multiband Image Denoising. *IEEE Trans. Image Process.* **2006**, *15* (3), 654–665.
- (20) Jansen, M.; Bultheel, A. Empirical Bayes Approach to Improve Wavelet Thresholding for Image Noise Reduction. *J. Am. Stat. Assoc.* **2001**, *96* (454),

629–639.

- (21) Liu, Y.; Cheng, X. A New Signal Denoising Algorithm from Wavelet Modulus Maxima. In *Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007)*; IEEE, 2007; pp 32–35.
- (22) Mallat, S. G. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **1989**, *11* (7), 674–693.
- (23) Meng, X.; He, Z.; Feng, G.; Xiao, B. An Improved Wavelet Denoising Algorithm for Wideband Radar Targets Detection. *Circuits, Syst. Signal Process.* **2013**, *32* (4), 2003–2026.
- (24) Fodor, I. K. Denoising through Wavelet Shrinkage: An Empirical Study. *J. Electron. Imaging* **2003**, *12* (1), 151.
- (25) Wink, A. M.; Roerdink, J. B. J. M. Denoising Functional MR Images: A Comparison of Wavelet Denoising and Gaussian Smoothing. *IEEE Trans. Med. Imaging* **2004**, *23* (3), 374–387.
- (26) Ernst, R. R. Sensitivity Enhancement in Magnetic Resonance. I. Analysis of the Method of Time Averaging. *Rev. Sci. Instrum.* **1965**, *36* (12), 1689.
- (27) Weil, J. E.; Bolton, J. R.; Wertz, J. E. *Electron Spin Resonance*; Wiley and Sons New York: Dordrecht, 1994.
- (28) Kak, A. C.; Slaney, M. *Principles of Computerized Tomographic Imaging*; IEEE Press: New York, 1988.
- (29) Donoho, D. L. De-Noising by Soft-Thresholding. *IEEE Trans. Inf. Theory* **1995**, *41* (3), 613–627.
- (30) Bryant, J. H. Electron Paramagnetic Resonance Image Reconstruction for Arbitrary Projection Sampling and Regions of Interest, The University of Chicago, 2011.
- (31) Donoho, D. L.; Johnstone, J. M. Ideal Spatial Adaptation by Wavelet Shrinkage. *Biometrika* **1994**, *81* (3), 425–455.
- (32) Johnstone, I. M.; Silverman, B. W. Wavelet Threshold Estimators for Data with

Correlated Noise. *J. R. Stat. Soc. Ser. B* **1997**, 59 (2), 319–351.

- (33) Chang, F.; Hong, W.; Zhang, T.; Jing, J.; Liu, X. Research on Wavelet Denoising for Pulse Signal Based on Improved Wavelet Thresholding. In *2010 First International Conference on Pervasive Computing, Signal Processing and Applications*; IEEE, 2010; pp 564–567.
- (34) Stein, S. M. Estimation of the Mean of a Multivariate Normal Distribution. *Ann. Stat.* **1981**, 9 (6), 1135–1151.
- (35) Donoho, D. L.; Johnstone, I. M. Adapting to Unknown Smoothness via Wavelet Shrinkage. *J. Am. Stat. Assoc.* **1995**, 90 (432), 1200.
- (36) Gao, H.-Y. Wavelet Shrinkage Denoising Using the Non-Negative Garrote. *J. Comput. Graph. Stat.* **1998**, 7 (4), 469–488.
- (37) Lin, Y.; Cai, J. A New Threshold Function for Signal Denoising Based on Wavelet Transform. In *2010 International Conference on Measuring Technology and Mechatronics Automation*; IEEE, 2010; pp 200–203.
- (38) Poornachandra, S.; Kumaravel, N.; Saravanan, T. K.; Somaskandan, R. WaveShrink Using Modified Hyper-Shrinkage Function. In *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*; IEEE, 2005; pp 30–32.
- (39) Lei Zhang; Paul Bao. Denoising by Spatial Correlation Thresholding. *IEEE Trans. Circuits Syst. Video Technol.* **2003**, 13 (6), 535–538.
- (40) Freed, J. H. New Technologies in Electron Spin Resonance. *Annu. Rev. Phys. Chem.* **2000**, 51 (1), 655–689.
- (41) Zhang, F.; Ma, L. MRI Denoising Using the Anisotropic Coupled Diffusion Equations. In *2010 3rd International Conference on Biomedical Engineering and Informatics*; IEEE, 2010; Vol. 1, pp 397–401.
- (42) Hofbauer, W.; Earle, K. A.; Dunnam, C. R.; Moscicki, J. K.; Freed, J. H. High-Power 95 GHz Pulsed Electron Spin Resonance Spectrometer. *Rev. Sci. Instrum.* **2004**, 75 (5), 1194–1208.
- (43) Freed, J. H. ESR and Molecular Dynamics. In *Biomedical EPR, Part B:*

*Methodology, Instrumentation, and Dynamics*; Kluwer Academic Publishers-Plenum Publishers: New York; pp 239–268.

- (44) Poole, C. P. *Electron Spin Resonance: A Comprehensive Treatise on Experimental Techniques*; Dover Publications, 1983.
- (45) Earle, K. A.; Dzikovski, B.; Hofbauer, W.; Moscicki, J. K.; Freed, J. H. High-Frequency ESR at ACERT. *Magn. Reson. Chem.* **2005**, *43*, S256-66.
- (46) Wang, Z.; Bovik, A.; Sheikh, H. R.; Simoncelli, E. P. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Trans. Image Process.* **2004**, *13* (4), 600–612.
- (47) Cohen, R. *Signal Denoising Using Wavelets*; 2012.
- (48) Dzikovski, B. G.; Livshits, V. A. EPR Spin Probe Study of Molecular Ordering and Dynamics in Monolayers at Oil/water Interfaces. *Phys. Chem. Chem. Phys.* **2003**, *5* (23), 5271.
- (49) Pavia, D.; Lampman, G.; Kriz, G.; Vyvyan, J. *Introduction to Spectroscopy*; Cengage Learning, 2008.
- (50) Galloway, C. M.; Le Ru, E. C.; Etchegoin, P. G. An Iterative Algorithm for Background Removal in Spectroscopy by Wavelet Transforms. *Appl. Spectrosc.* **2009**, *63* (12), 1370–1376.

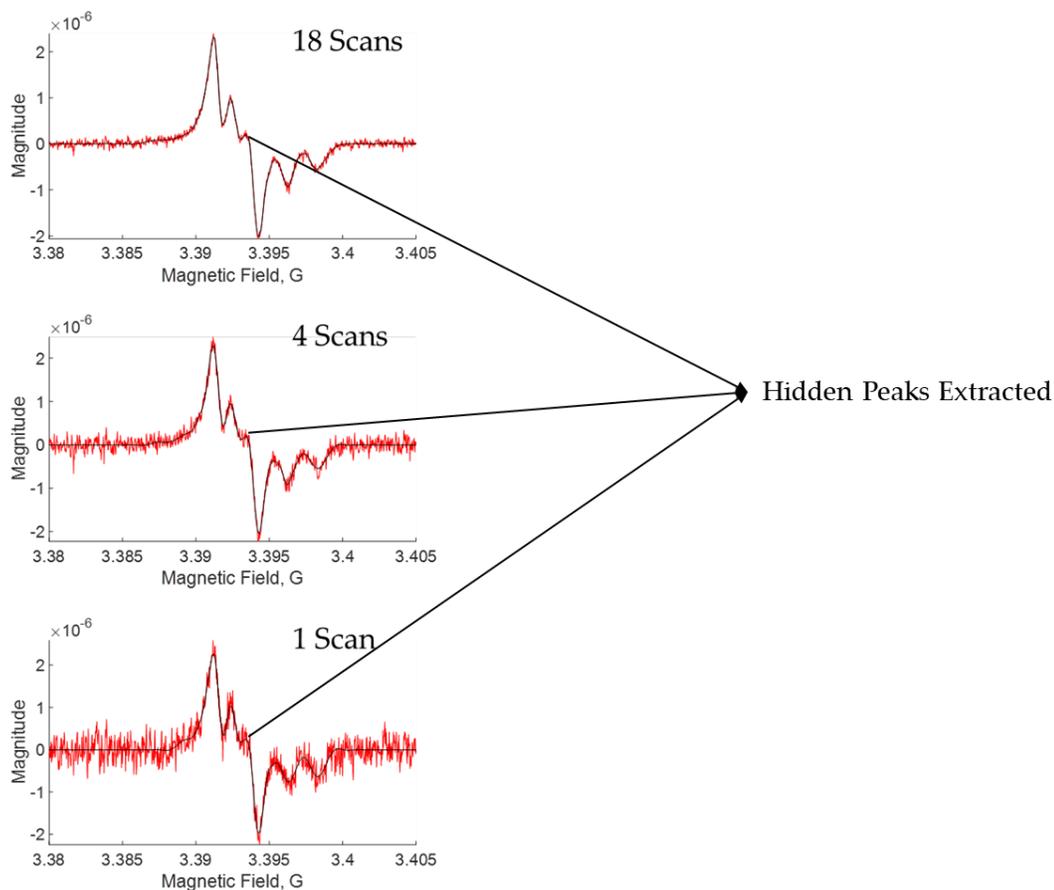
## CHAPTER 3

### NOISE ELIMINATION AND REDUCTION VIA DENOISING: A WAVELET APPROACH FOR 1-D SPECTROSCOPY

#### Abstract

Experimental noise in one-dimensional spectroscopy impedes studies by reducing signal resolution and/or suppressing weak signals. Signal averaging and filtering are the primary methods used to reduce the noise, but they have limited effectiveness. We use a wavelet transform-based approach to effectively remove the signals. We apply the wavelet shrinkage method with newly developed methods on selecting decomposition levels, noise thresholds as well as noise thresholding the Approximation component in wavelet domain that contains low frequency information of the experimental signal. In addition, we use undecimated discrete wavelet transform, instead of decimated discrete wavelet transform, to improve signal resolution in wavelet domain. The denoised results from many types of 1D signals show that the wavelet denoising is accurately reproduces the signal peaks, including weak signals. Compared to decimated discrete wavelet transform, the utilization of undecimated discrete wavelet transform significantly improves the signal-to-noise ratio (SNR). The method can be applied to any 1D spectroscopic signals.

## Graphic Abstract



\* This chapter is in preparation for a journal publication

**Authors:** Madhur Srivastava, Boris Dzikovski, and Jack H. Freed

**Author Contributions:** Madhur Srivastava conducted the research and wrote the paper, Boris Dzikovski collected the experimental data and generated the model data code, and Jack Freed supervised the research and edited the paper.

### 3.1 Introduction

Spectroscopic methods are widely used in the fields of chemistry and biophysics, among others, to reveal underlying mechanisms and properties that are key to the understanding of the chemical and biophysical systems.<sup>1</sup> They play an important role to study structure and function of molecules and biomolecules and their interactions in various environments. Nuclear Magnetic Resonance, Electron Spin Resonance, Infrared, Optical and Raman spectroscopy are some of the commonly used spectroscopic methods, among others, with a variety of applications.<sup>1</sup>

Despite major advances in their technology, they frequently still lack sufficient sensitivity and resolution to conduct experiments in the desired environment. They usually operate with limited efficacy and can only probe samples in a constrained environment. This is mainly due to the presence of noise in the signals obtained from the instruments.

Signal averaging is the primary method to reduce noise but is time consuming. It reduces the noise by a factor of  $\sqrt{N}$ , where  $N$  is the number of repeats or scans of a sample.<sup>2</sup> With signal averaging, long signal collection time is frequently needed to obtain a sufficient Signal-to-Noise Ratio (*SNR*). This is not suitable for unstable samples as well as samples with low concentration. Weak signals may well be suppressed under noise.

Many data processing methods such as noise filtering and model fitting can be used to remove noise. Noise filtering methods<sup>3,4</sup> (such as Gaussian filtering, low pass filtering, mean and median filtering, and anisotropic filtering, to name a few) have had limited success and are prone to signal distortion and incomplete noise removal. They mostly remove high frequencies, possibly including those of signal, and leave low frequencies of noise unaffected. On the other hand, model fitting methods require *a priori* information about the signal, which is not possible in many experimental cases.

We have developed a wavelet transform-based signal denoising method for spectroscopic signal signals that can successfully identify and separate noise from signal with *SNR* as low as 3 (cf. Chapter 2). We call our method “NERD” (Noise Elimination and Reduction via Denoising). NERD can be applied to any spectroscopic method as it does not require *a priori* or specialized information about the signal to carry out denoising. It utilizes the property of wavelet transforms that allows simultaneous representation of the signal domain, typically the time domain, and its frequency information. For example, one cannot access frequency information of a signal in the signal (e.g. time) domain. Similarly, signal domain information of a signal is lacking in the Fourier-transformed frequency domain. But signal and noise have non-overlapping and different characteristics and features, which a wavelet transform is able to exploit. Noise and signal usually do not have the same frequency at the same signal location which can be easily be identified if signal domain and frequency information are simultaneously available. Thus, while filtering methods can only either keep a frequency component or remove it, a wavelet transform can

distinguish noise and signal with the same frequency and remove the noise by using their different signal domain location.

Wavelet transform-based denoising methods<sup>5-20</sup> have been around for 25 years but have failed to achieve their true potential in the past. Despite the fact that noise and signal can be distinguished and separated in the wavelet domain, *standard* wavelet denoising methods<sup>21,22,3</sup> are unable to determine the needed parameters to achieve that. However, recent developments in obtaining desired object parameters such as noise thresholds and identifying wavelet components with noise presence have been shown to successfully denoise experimental signals at low *SNR* (cf. Chapter 4).<sup>23</sup> One such application has been shown in the experimental time domain signals of the electron spin resonance based pulsed dipolar spectroscopy.<sup>24</sup> The denoising method called WavPDS<sup>24</sup> was able to retrieve signal from the *SNR* as low as 3 that led to revealing of multiple conformations in biomolecules which are concealed due to the presence of instrumental noise.

We have adapted our recently developed wavelet denoising method with some modifications for denoising spectroscopic signals. We show the effectiveness of NERD by testing it on continuous wave-electron spin resonance spectroscopy<sup>25</sup> (cw-ESR) data. Both model and experimental data are used and different cases are examined. NERD is successfully able to retrieve signal from noise, including weak signals as well as overlapping signals. We show that the signal resolution is greatly improved with a two order-of-magnitude increase in *SNR*. Further, with fewer scans

needed to reproduce the original signal, the signal acquisition time can be reduced by at least an order-of-magnitude. The method is general and can be applied to any one-dimensional spectroscopy data.

We note that in cw-ESR the signal domain is achieved by varying the external magnetic field,  $B_0 - B_0(t)$ .

## 3.2 Methods and Experiments

### 3.2.1 Wavelet Transforms and Discrete Wavelet Transforms

A signal can be analyzed in the signal domain (such as time, space, and others) where it is acquired or a transformed domain in which some characteristics/features of a signal are especially prominent.<sup>26</sup> The Fourier transform (FT) is the most common as it provides frequency information for the complete signal. However, the FT cannot provide frequency information for a “localized” signal region. It also does not show how frequency information is changing from the start of the signal to its end. Wavelet transforms can provide both these types of information simultaneously. Furthermore, in a wavelet transform, signal localization varies from single data points to the complete signal. Wavelet transforms also reveal how a signal looks when observed from a specific frequency or frequency range. These all allow comprehensive analysis of a signal from multiple parameters and information sets. Mathematically, a wavelet transform is defined as<sup>27</sup> (cf. Chapter 1),

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

(3.1)

where  $s$  is the inverse frequency (or frequency range) parameter,  $\tau$  is the signal localization 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 called “wavelet.” Different wavelets are used to vary selectivity or sensitivity of adjacent frequencies with respect to signal localization. They are not dependent on *a priori* information of the signal or its characteristics. Depending on the application, different wavelets can be selected.

Like the FT, a wavelet transform has high computational complexity and hence, for practical implementation, the discrete wavelet transform (DWT) is used. To avoid redundancy and reduce computational complexity, wavelet transforms are taken over non-overlapping frequency ranges called “sub-bands” or “Detail components” in wavelet terminology. To differentiate between different Detail components, sub-bands containing the highest frequencies are called “Detail component at decomposition level 1” and sub-bands containing lower frequencies are labeled in ascending numerical order of decomposition level. At each decomposition level, the residual lower frequency sub-band is stored as the “Approximation component” at that all decomposition level, so that signal information at all frequencies is contained at any given decomposition level. For a discrete signal with length  $p$ , the maximum number of decomposition levels that can be obtained is  $N$ , where  $N = \log_2 p$ . The decomposition level can be referred to as  $j$ , where  $(1 \leq j \leq N)$ .

For DWT, the Detail and Approximation components are defined as<sup>27</sup>,

$$D_j[n] = \sum_{m=0}^{p-1} f[t_m] 2^{j/2} \psi[2^j t_m - n] \quad (3.2)$$

and

$$A_j[n] = \sum_{m=0}^{p-1} f[t_m] 2^{j/2} \phi[2^j t_m - n] \quad (3.3)$$

where  $f[t_m]$  is the discrete input signal (analogous to  $f(t)$ ),  $p$  is the length of input signal  $f[t_m]$ ,  $D_j[n]$  and  $A_j[n]$  are the Detail and Approximation components, respectively, at  $j^{th}$  decomposition level, and  $\psi[2^j t_m - n]$  and  $\phi[2^j t_m - n]$  are wavelet and scaling functions, respectively. The scaling and wavelet functions, at a decomposition level, are orthogonal to each other as they represent non-overlapping frequency information. Similarly, wavelet functions at different decomposition levels are orthogonal to each other.

The signal  $f[t_m]$  can be reconstructed using the Inverse DWT (IDWT) as follows,

$$f[t_m] = \sum_{k=0}^{p-1} A_{j_0}[k] \phi_{j_0,k}[t_m] + \sum_{j=1}^{j_0} \sum_{k=0}^{p-1} D_j[k] \psi_{j,k}[t_m] \quad (3.4)$$

where  $j_0$  is a given decomposition level from which input signal needs to be reconstructed.

For more information on Wavelet Transform and DWT, refer to Chapter 1 (Signal Analysis and Denoising: Limitations of Fourier Transform and Short Time Fourier Transform, and Introduction to Wavelet Transform).

### **3.2.2 Wavelet Denoising**

The wavelet denoising procedure has three major steps.<sup>28</sup> First, the DWT of the input signal at  $k^{th}$  decomposition level is taken. Second, Detail components 1 to  $k$  and  $k^{th}$  Approximation component are noise thresholded. Third, IDWT of the noise thresholded components are taken. The  $k^{th}$  Approximation component and  $k$  Detail components together constitute the complete input signal as can be seen from equation 3.4. Following are the steps used in NERD to accomplish denoising.

#### ***3.2.2.1 Undecimated Discrete Wavelet Transform***

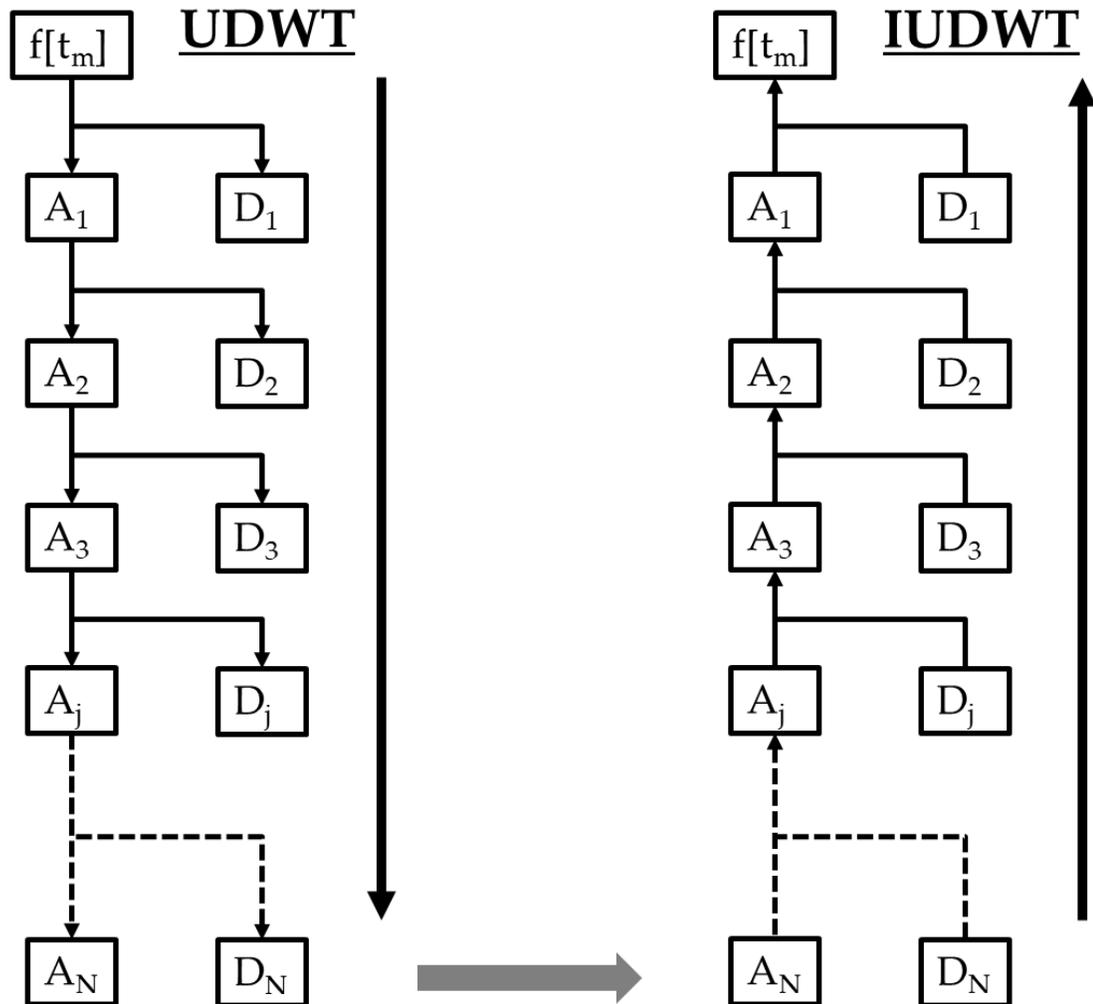
In NERD, we use undecimated DWT<sup>29</sup> to achieve the maximum signal and noise resolution possible in wavelet transform (cf. Fig 3.1). This means that each Detail and Approximation component has the same length as of the input signal. For instance, a signal with 128 data length will have 7 Detail and Approximation components (from 7 decomposition levels), each having the length of 128 data points. Using undecimated

DWT improves the resolution between adjacent data points (and also noise and signal) as their information is represented separately by wavelet coefficients.

In the standard procedure, decimated DWT is used which downsamples the Detail and Approximation components by 2 at each decomposition level, i.e. alternate data points are removed (cf. Fig 3.2). It preserves the number of input signal and wavelet coefficients. For example, a signal with a length of 128 data points will have 64, 32, 16, 8, 4, 2 and 1 data points for both Detail and Approximation components at decomposition level 1, 2, 3, 4, 5, 6 and 7, respectively. This downsampling does not affect the accurate signal reconstruction because it satisfies the “Nyquist rate” condition (the minimum number of data points required to perfectly reconstruct the signal).

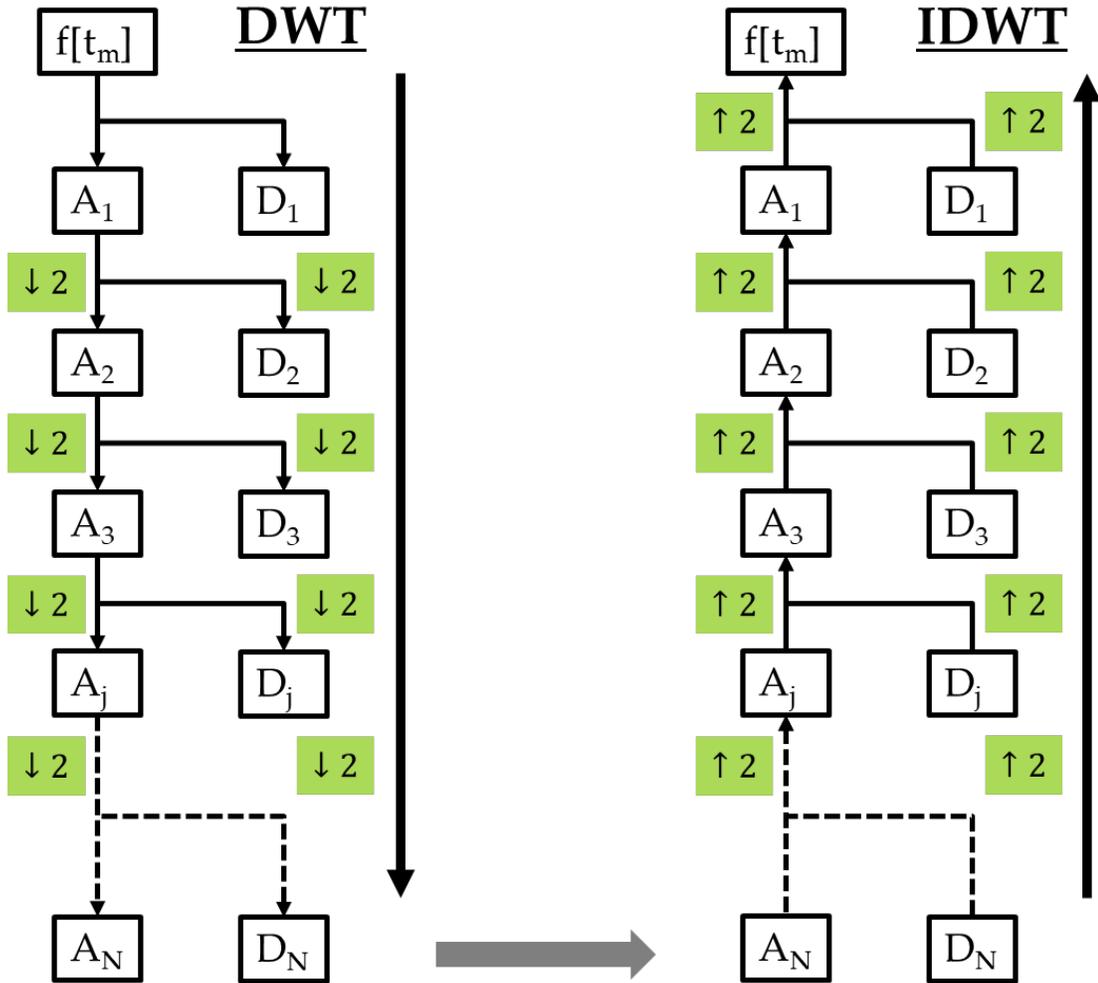
However, the reduced number of wavelet coefficients also reduces the information available to distinguish noise coefficients from signal coefficients. With downsampling of data by 2 at each stage of decomposition level, the signal resolution is reduced by factor of 2. The selectivity and sensitivity of frequency information (called resolution) of a signal are blurred, especially at lower frequencies. Removing noise in experimental signals requires high resolution at all frequencies.

## Undecimated DWT and IDWT



**Figure 3.1:** The block diagram of the undecimated discrete wavelet transform (DWT) and inverse discrete wavelet transform (IDWT).  $f[t_m]$  is the discrete input signal, and  $D_j$  and  $A_j$  are the Detail and Approximation components, respectively, at decomposition level  $j$ .

## Decimated DWT and IDWT



**Figure 3.2:** The block diagram of the undecimated discrete wavelet transform (DWT) and inverse discrete wavelet transform (IDWT).  $f[t_m]$  is the discrete input signal, and  $D_j$  and  $A_j$  are the Detail and Approximation components, respectively, at decomposition level  $j$ . The symbols  $\downarrow 2$  and  $\uparrow 2$  represent downsampling and upsampling by a factor of 2, respectively.

### ***3.2.2.2 Wavelet Selection***

There are many standard wavelet families that can be used for the DWT. We use the coiflet family wavelet “coiflet-3”. The coiflet family satisfies the mini-max condition which minimizes the error in extracting the local features. This feature is essential for enhancing signal resolution between nearby data points as well as for weak signals. Coiflet-3 wavelet is selected for its appropriate length. A smaller length may not capture all the necessary information, whereas a larger length would yield redundant information.

### ***3.2.2.3 Decomposition Level Selection***

For successful noise thresholding in the wavelet domain, all the Detail components and the highest decomposition Approximation component that still contains noise needs to be identified. This is accomplished by selecting the maximum decomposition level  $k$  where noise is still present keeping Detail components from 1 to  $k$  and the  $k^{th}$  Approximation component. Inaccurate selections of  $k$  result in either incomplete noise removal or signal distortion. To identify if a Detail component contains noise, we first calculate the “sparsity” of each Detail component in the following way,

$$S_j = \frac{\max(D_j)}{\sum_{n=1}^{q_i} D_j[n]} \tag{3.5}$$

where  $S_j$  is the sparsity and  $q_i$  is the length of the Detail component. It can be seen that  $S_j$  measures the sparsity of the Detail component at decomposition level  $j$ .  $S_j$  will have smaller values for noisy Detail components compared to virtually noise-free Detail components. Noise presence reduces sparsity. We use a cut-off  $T_r$  to identify noise and noise-free Detail components. The maximum decomposition level  $k = j$  is for which  $S_j < T_r$  and  $S_{j+1} > T_r$ .

#### ***3.2.2.4 Noise Thresholds Calculation***

In the wavelet domain, each noisy Detail component contains noise and signal coefficients that are typically separate from each other. Wavelet coefficients representing random noise (generated from instruments) have many small magnitudes, whereas wavelet coefficients representing signal coefficients are fewer but have large magnitudes due to their coherent nature. Coherent signals can easily be captured using those few coefficients with high magnitudes. Differences in magnitude of signal (large) versus noise (small) enable us to use a threshold that can separate them. Selecting an accurate threshold is necessary to avoid incomplete noise removal and/or signal distortion. We calculate lower and upper noise thresholds for negative and positive wavelet coefficients in the Detail component. Detail components are always centered around 0 because they represent non-zero frequencies. The noise thresholds are calculated using the following formulae:

$$\lambda_{j,L} = \mu_j - \kappa_{j,L}\sigma_j$$

(3.6)

and,

$$\lambda_{j,H} = \mu_j - \kappa_{j,H}\sigma_j \quad (3.7)$$

where  $\lambda_{j,L}$  and  $\lambda_{j,H}$  are the lower and upper thresholds,  $\mu_j$  and  $\sigma_j$  are the mean and standard deviation of the Detail component, and  $\kappa_{j,L}$  and  $\kappa_{j,H}$  are adjustable parameters associated with the lower and upper thresholds, respectively.  $j$  refers to the decomposition level of the Detail component and satisfies  $1 \leq j \leq k$ . The  $\kappa_{j,L}$  and  $\kappa_{j,H}$  values can be obtained from the formulae presented in Chapter 2.

### ***3.2.2.5 Noise Thresholding Function***

The noise thresholds are applied in the Detail component using the hard thresholding function as follows,

$$D'_j[n] = \begin{cases} 0, & \text{for } \lambda_{j,L} < D_j[n] < \lambda_{j,H} \\ D_j[n], & \text{otherwise} \end{cases} \quad (3.8)$$

where  $D'_j[n]$  is the denoised (or noise-thresholded) Detail component. Hard thresholding is used because each wavelet coefficient is either noise coefficient or signal coefficient (cf. Chapter 2). Applying soft thresholding would result in signal distortion.

### 3.2.2.6 Denoising Approximation Component

Unlike standard denoising methods, in NERD noise thresholding is also applied to the Approximation component (as shown in the Chapter 2). The noise thresholds are applied in the  $k^{th}$  Approximation component and are calculated similar to the Detail components using the equations 3.6 and 3.7. However, the Approximation component may not be centered around 0, as it contains a d.c. frequency (i.e., frequency at  $0^{th}$  value). Therefore, its lower and upper thresholds can have any value and do not need to represent negative and positive wavelet coefficient values, respectively. The hard noise thresholding in the equation 3.8 is modified for the Approximation component in the following way:

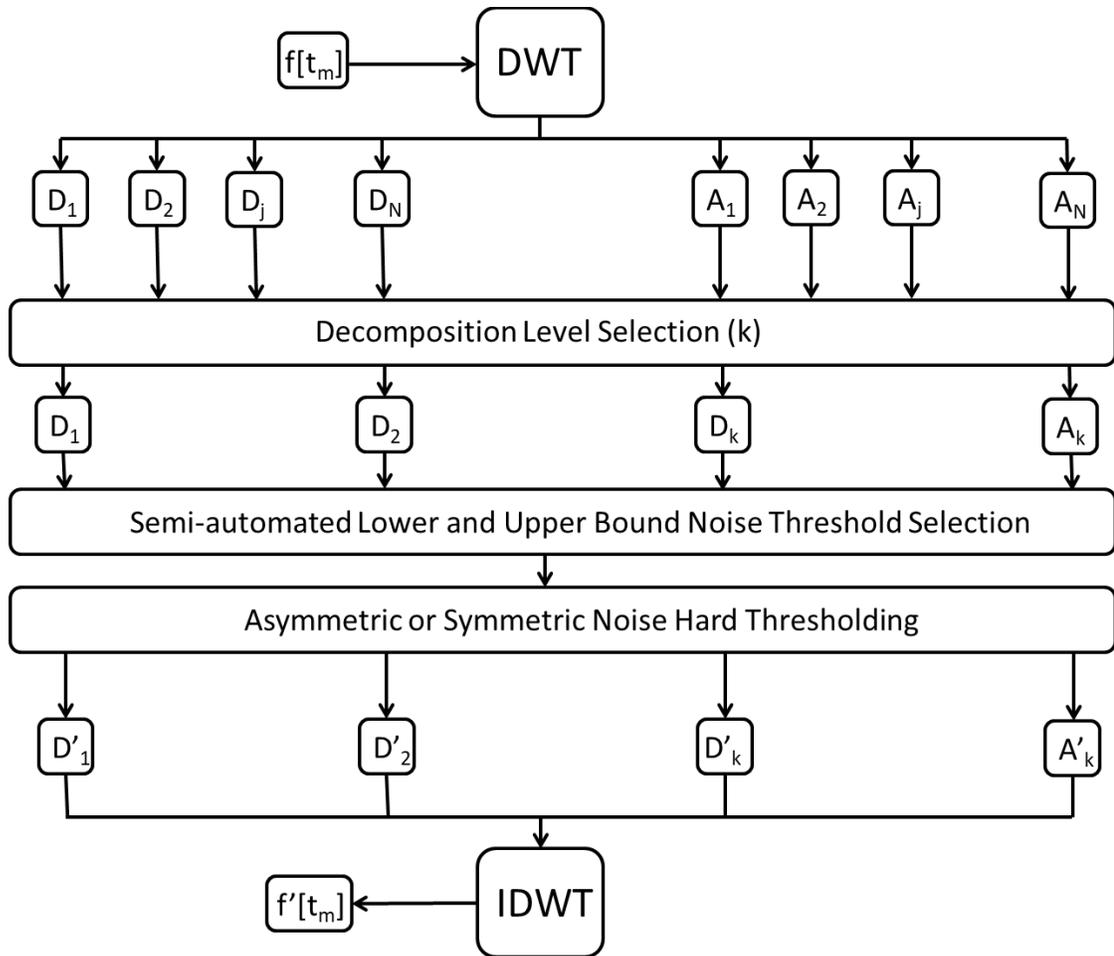
$$A'_k[n] = \begin{cases} a, & \text{for } \lambda_{j,L} < A_k[n] < \lambda_{j,H} \\ A_k[n], & \text{otherwise} \end{cases} \quad (3.9)$$

where  $A_k[n]$  and  $A'_k[n]$  are the noisy and denoised Approximation components, respectively, at the  $k^{th}$  decomposition level, and  $a$  is the constant with which noisy values is replaced. For signal with baseline centered around 0,  $a = 0$ . For non-zero baseline signal,  $a$  is assigned the baseline value.

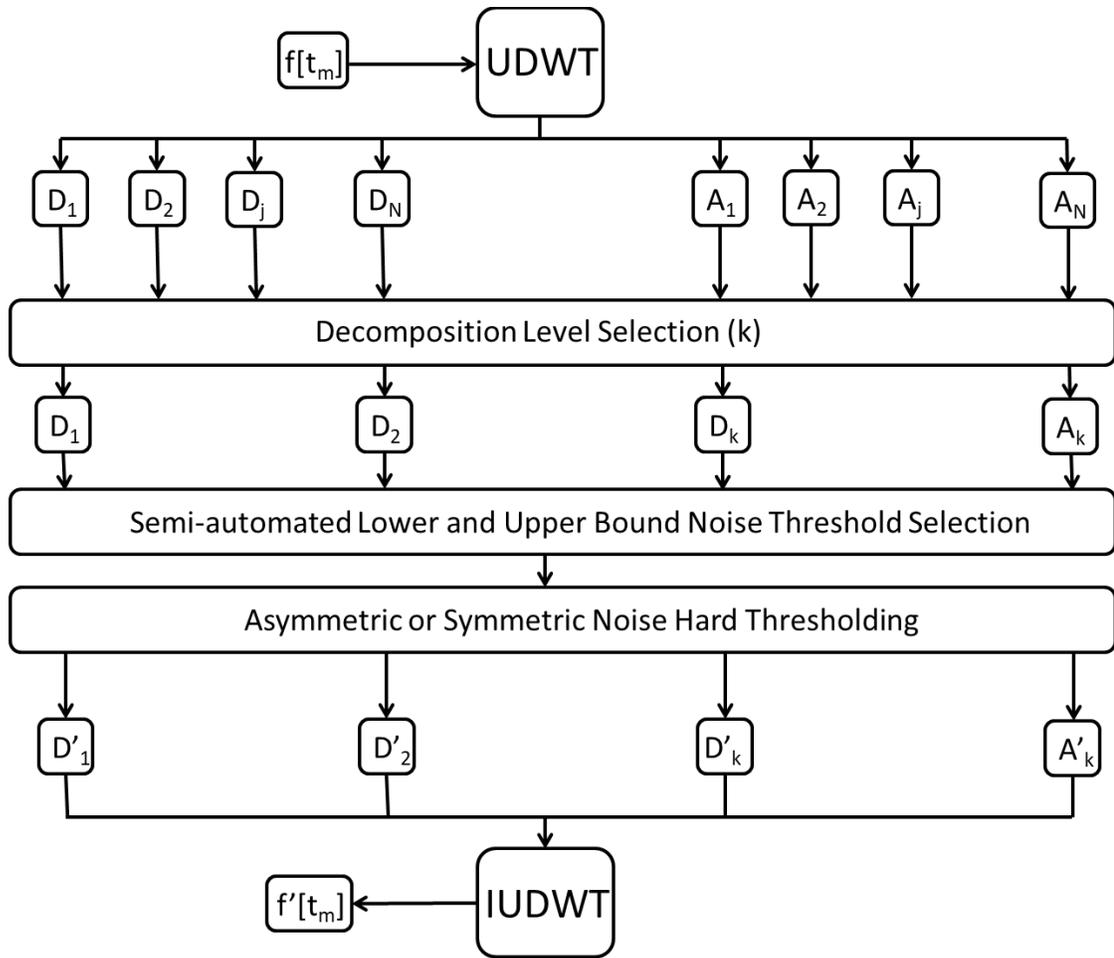
### 3.2.3 Difference between NERD and Standard Denoising

The NERD uses our recently developed new wavelet denoising method (cf. Chapter 2) to select decomposition level, calculate noise thresholds and apply hard noise thresholding based in lower and upper thresholds (cf. Fig. 3.3). The noise thresholds are applied to both the Detail and Approximation component (cf. Fig. 3.3). The NERD improves the new wavelet denoising method by employing undecimated DWT (cf. Fig. 3.1), instead of decimated DWT (cf. Fig. 3.2). This results in enhanced signal resolution in the wavelet domain with a reduced the minimum SNR at and above which the experimental signal can be denoised.

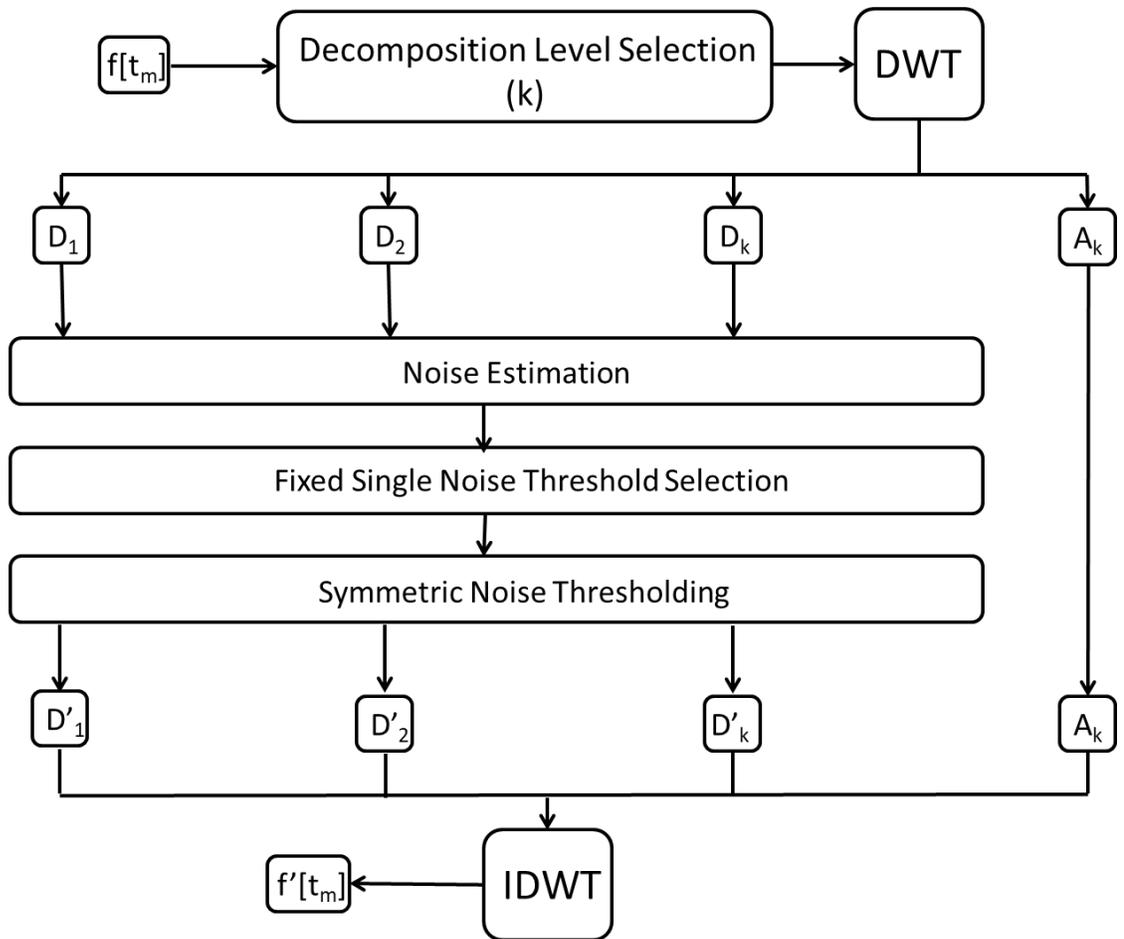
The standard denoising methods, however, rely on an arbitrary value of  $k$  to select the maximum decomposition levels to denoise. They utilize a single noise threshold for the magnitudes of negative and positive Detail components, making it vulnerable to incomplete noise removal and/or signal distortion, especially if noise and signal have asymmetric characteristics. This also limits the noise thresholding to the Detail components because the Approximation component may not be centered around 0. Noise estimation methods used to obtain noise thresholds are not definitive. Lastly, they use the decimated DWT. Fig. 3.5 shows the standard wavelet denoising procedure.



**Figure 3.3:** The block diagram of the new wavelet denoising method.  $f[t_m]$  and  $f'[t_m]$  are the noisy and denoised signals, respectively, and  $D_j$  and  $A_j$  are the Detail and Approximation components, respectively, at decomposition level  $j$ . The DWT and IDWT stands for Discrete Wavelet Transform and Inverse Discrete Wavelet Transform, respectively.



**Figure 3.4:** The block diagram of the Noise Elimination and Reduction via Denoising (NERD) method.  $f[t_m]$  and  $f'[t_m]$  are the noisy and denoised signals, respectively, and  $D_j$  and  $A_j$  are the Detail and Approximation components, respectively, at decomposition level  $j$ . The DWT and IDWT stands for Discrete Wavelet Transform and Inverse Discrete Wavelet Transform, respectively.



**Figure 3.5:** The block diagram of the standard wavelet denoising methods.  $f[t_m]$  and  $f'[t_m]$  are the noisy and denoised signals, respectively, and  $D_j$  and  $A_j$  are the Detail and Approximation components, respectively, at decomposition level  $j$ . The DWT and IDWT stands for Discrete Wavelet Transform and Inverse Discrete Wavelet Transform, respectively.

### 3.2.4 Spectroscopy Method

We use cw-ESR spectroscopy model and experimental data for testing the NERD method. It is extensively applied to study the dynamics and structure of biomolecules and is the most commonly used ESR technique.<sup>25</sup> The cw-ESR spectra can be constituted from Lorentzian functions, making it easy to generate model data that replicates experimental signals. The cw-ESR spectrum is acquired in magnetic field ( $B_0$ ) domain, which is varied, i.e.  $B_0 = B_0(t)$ . Moreover, both in model and experimental data, a variety of spectra can be generated as test signals to the NERD denoising methods.

### 3.2.5 Model Data Generation

The model data is generated using 3 spectral lines obtained as the derivatives of Lorentzian functions. Two of the spectral lines are weak signals whose peak height is 3% of the third, which is the main signal. Gaussian noise is added to the model data to generate *SNRs* of 30 and 10 w.r.t the peak height of the main signal.

### 3.2.6 Experimental Data Generation

#### 3.2.6.1 ESR Experiment and Sample Preparation

Two experimental signals were acquired as examples to test the method. Example 1 contains three main lineshapes and 6 weak signals. The sample used was Tempol (4-Hydroxy-2,2,6,6-Tetramethylpiperidine 1-oxyl).<sup>30-34</sup> which is a commonly used spin-

probe molecule. The sample concentration was 4  $\mu\text{L}$  in a 100  $\mu\text{M}$  aqueous solution. The ESR experiment was performed at 20°C on a commercial spectrometer (*BRUKER ELEXYS-II E500*) at X-band (9.4GHz) microwave frequency that corresponds to 0.34 Tesla dc magnetic field. The spectral data was of 4096 length along the magnetic field sweep.

Example 2 contains multiple overlapping spectrum, providing a complex signal to study. The sample consisted of 5  $\mu\text{L}$  of phospholipid vesicles doped with 0.5% of a lipid spin label: 16-PC (1-acyl-2-[16-(4,4-dimethyloxazolidine-N-oxyl)stearoyl]-sn-glycero-3-phosphocholine) in the fluid phase. It was then suspended in water. A home-built 95 GHz ESR spectrometer with a dc magnetic field of 3.3 Tesla, developed at ACERT, was used to collect the experimental data at 25°C. The spectral data was of 512 data length.

### ***3.2.6.2 Signal Averaging***

In example 1, two noisy signals were generated by averaging 4 and 16 scans that resulted in an *SNR* of 14 and 30, respectively, for the main signal and an *SNR* of 1.5 and 0.5, respectively, for the weak signal (cf. Table 2). A reference signal was also generated by averaging 500 scans. For example 2, three noisy signals were generated by averaging 1, 4, and 18 scans. Their respective *SNRs* were 10, 19, and 39.

### 3.2.7 Objective Measure

To test the efficacy of the NERD method, the Signal-to-Noise Ratio (*SNR*) is used for calculating the noise presence in the noisy and denoised signals. The *SNR* is defined as,

$$SNR = \frac{\max(|f[t_m]|)}{RMS(Noise)} \quad (3.10)$$

where  $RMS(Noise)$  is the root-mean-square of the noise present in the signal. The noise presence is measured from the spectral positions where there is no signal present in the model and experimental data. Note that the *SNR* is the inverse of the amount of noise present with respect to the maximum peak height of the signal. This implies that the remaining signals, whose maxima are less than the maximum peak height, have lower *SNR*.

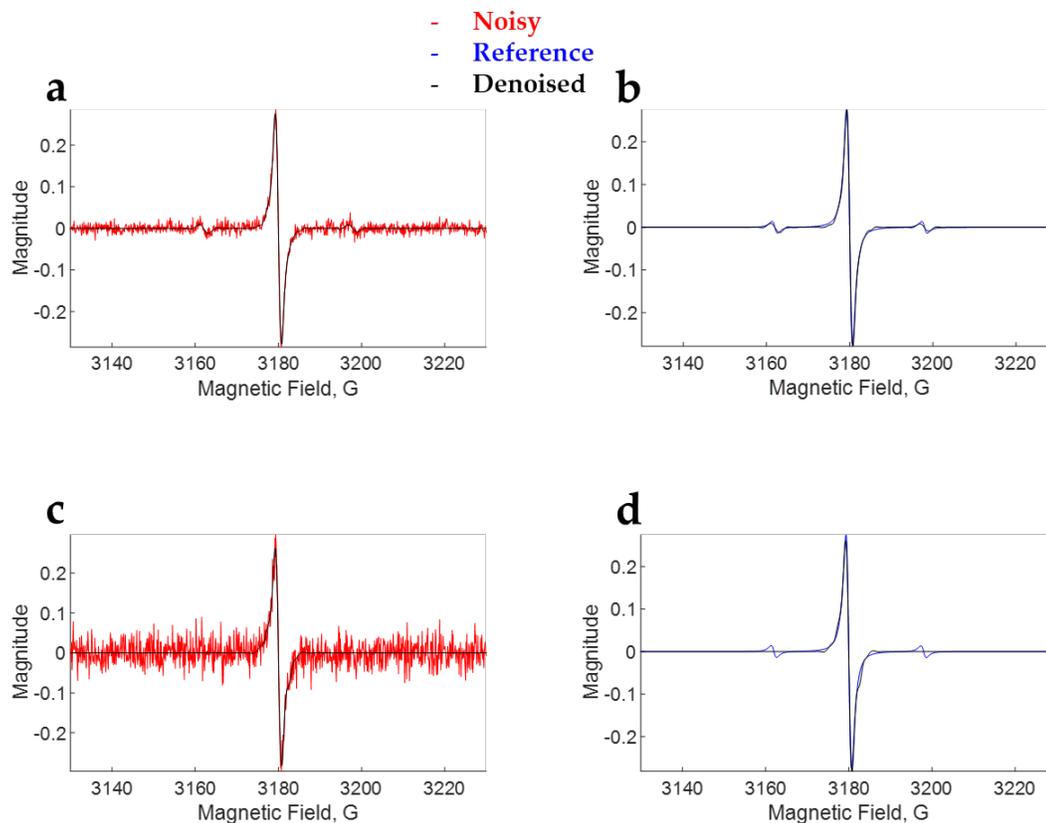
## 3.3 Results

### 3.3.1 Model Data

In the Fig. 3.6, noisy and denoised data are overlapped in subfigures a and c for *SNRs* 30 and 10, respectively. The denoised data is then compared to the noise-free model data in the subfigures b and d. It can be seen that for *SNR* of 30 (cf. subfig. a and b), NERD perfectly reproduces the noise-free signal. Even the weak signals with *SNR* of 3 are accurately identified and retrieved by the denoising method. These weak signals were buried under the noise and are not clearly visible by eye. Table 1 demonstrates

the effectiveness of NERD. The *SNRs* of the denoised signal is  $4.33 \times 10^6$  and  $1.91 \times 10^5$  for main and weak signal, respectively, which is several orders-of-magnitude improvement compared to the noisy signal.

For *SNR* of 10, NERD effectively reproduces the main signal yielding an *SNR* of 2,212, which is more than two orders-of-magnitude greater than the noisy signal. However, the weak signals that have *SNR* of 0.5 were not retrieved by the method, showing its limitation for such low *SNR*. It shows that if the noise magnitude is greater than the signal, NERD may not be able to separate noise.



**Figure 3.6:** Model Data – Comparison of noisy signal and denoised signal obtained by NERD at SNRs 30 and 10. a) Noisy signal at peak height SNR of 30 overlapped with its denoised signal, b) denoised signal at SNR 30 overlapped with the noise-free signal, c) Noisy signal at peak height SNR of 10 overlapped with its denoised signal, and d) denoised signal at SNR 10 overlapped with the noise-free signal.

**Table 3.1:** Model Data -- SNR of Noisy and Denoised Signals at Different SNRs

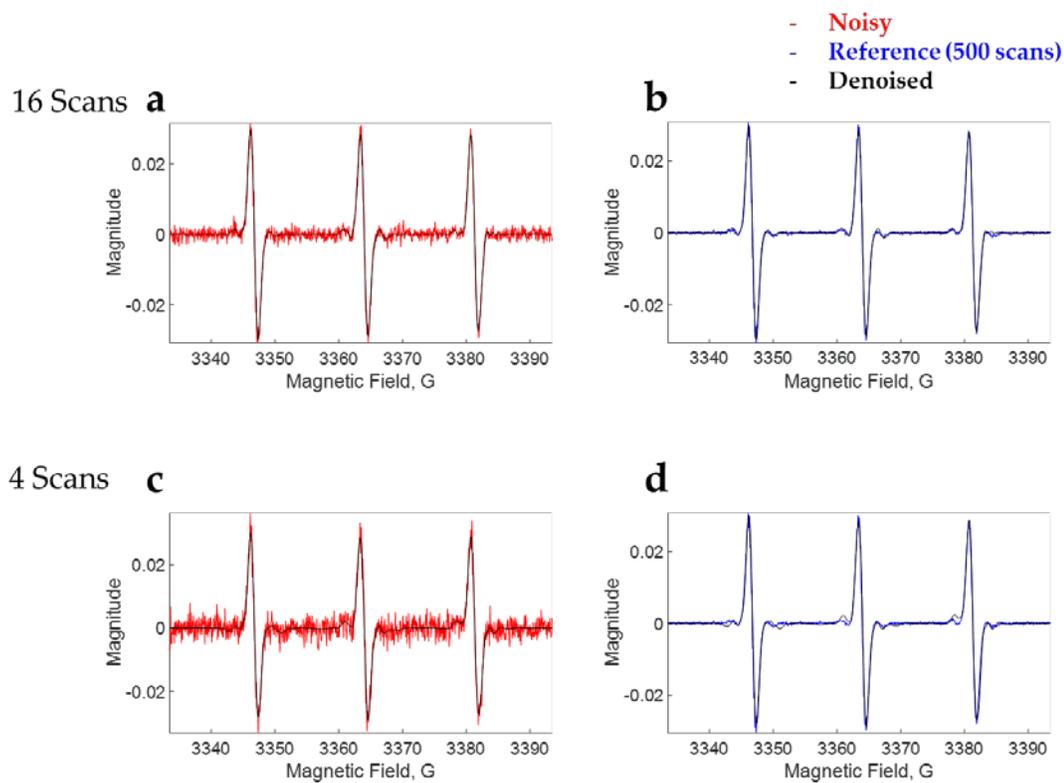
<b>Data</b>		<b>Noisy SNR</b>	<b>Denoised SNR</b>
Case 1	Main Signal	30	$4.36 \times 10^6$
	Weak Signal	3	$1.91 \times 10^5$
Case 2	Main Signal	10	2,212
	Weak Signal	0.5	N/A

### 3.3.2 Experimental Data

#### 3.3.2.1 Example 1: Non-overlapping signals

The result of example 1 is shown in Fig. 3.7. It can be seen that NERD is successfully able to remove noise from the 16-scan noisy data and reconstructs the desired signal with near perfection (cf. subfig. b). More importantly, the method is able to retrieve weak signals that are not visible in the noisy data. The *SNR* of the main peaks are increased from 30 to 46,416 and weak signals from 1.5 to 1613 (cf. Table 2), both of which are more than two orders-of-magnitude improvement in signal resolution. Note that there is no signal distortion.

Similarly, NERD also improves the signal resolution for the 4-scan data by more than two-orders of magnitude for both main peaks and weak signals. The *SNRs* are increased from 14 to 2,435 for the main peak and 0.5 to 85 for the weak signals after denoising. When compared to the reference signal, one can see that weak signals in denoised data for 4-scan are not as good as in 16 scan as expected, because of the intense noise amplitude where noise is twice the signal value. Mere retrieval of these weak signals is significant in itself as it informs about their presence. Recall that for the model data, denoising was not able to retrieve such signal.



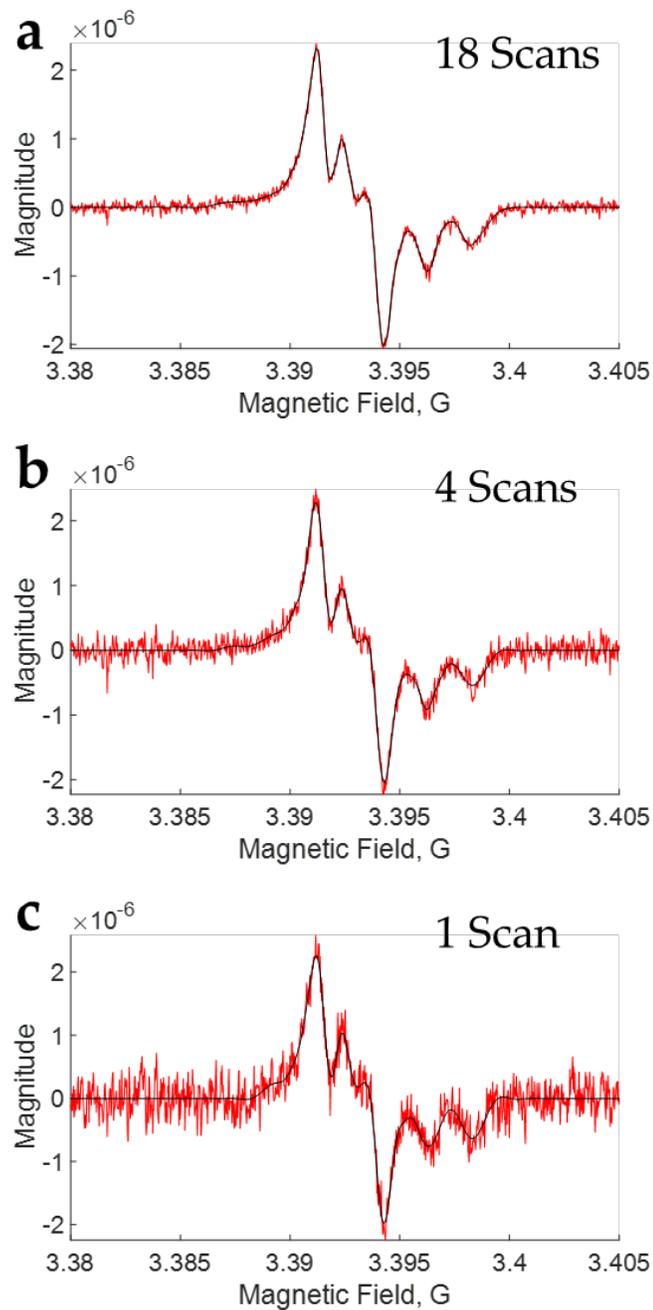
**Figure 3.7:** Experimental Data Example 1 – Results of NERD at 16 and 4 scans cw-ESR data with SNR of 30 and 14, respectively. a) 16-scan noisy signal at peak height SNR of 30 overlapped with its denoised signal, b) denoised signal of 16 scan noisy signal overlapped with the reference (500 scan) signal, c) 4-scan noisy signal at peak height SNR of 10 overlapped with its denoised signal, and d) denoised signal of 4 scan noisy signal overlapped with the reference (500 scan) signal.

**Table 3.2:** Example 1 -- SNR of Noisy and Denoised Signals at 18 and 4 Scans

<b>Data Averaging</b>		<b>Noisy SNR</b>	<b>Denoised SNR</b>
16 Scans	Main Signal	30	46,416
	Weak Signal	1.5	1,613
4 Scans	Main Signal	14	2,435
	Weak Signal	0.5	85

### 3.3.2.2 Example 2: Overlapping signals

In example 2 shown in Fig. 3.8, NERD is applied to 18-scan, 4-scan and 1-scan signal averaged noisy data. The *SNR* for each of them is respectively 39, 19, and 10. Similar to the previous model and experimental cases, it can be seen that the denoised results enhance the *SNR* by more than an order-of-magnitude to 1243, 947 and 379, respectively, for 18, 4 and 1 scans (cf. Table 3). More importantly, the denoised data is able to successfully extract the lineshapes that are suppressed under the noise. For example, the third peak in the subfigs. b and c is overwhelmed with noise with little evidence of its presence. After denoising, this peak is now well resolved as can be seen from the comparison with the 18-scan noisy and denoised spectra, both of which clearly show the existence of that signal peak.



**Figure 3.8:** Experimental Data Example 2 – Results of NERD at 18, 4 and 1 scans cw-ESR data with SNR of 39, 19 and 10, respectively. a) 18-scan noisy signal overlapped with it denoised data, b) 4-scan noisy signal overlapped with it denoised data, and c) 1-scan noisy signal overlapped with it denoised data.

**Table 3.3:** Example 2 -- SNR of Noisy and Denoised Signals at 18, 4, and 1 Scans

<b>Data Averaging</b>	<b>Noisy SNR</b>	<b>Denoised SNR</b>
18 Scans	39	1243
4 Scans	19	947
1 Scan	10	379

### 3.4 Discussion

The model and experimental results show that the NERD method can improve the *SNR* by many fold and accurately reconstruct the spectra without distortion or artifacts. The noise is easily identified and removed in the wavelet domain by NERD, which was not possible with current standard methods. The results suggest that signal components with *SNR* of 3 and above can be perfectly reconstructed, while the *SNR* below that depends on the chance as shown in Fig 3.7. The *SNR* is improved by at least an order-of-magnitude and many hidden details are successfully extracted. In particular, weak signals can be retrieved along with other spectra. This is important because many times these weak signals reveal information that cannot be obtained elsewhere. For example, the C-13 splitting is in hard to detect in spectroscopy due to its weak presence (it is only 1/100<sup>th</sup> of the abundant C-12 signal). Retrieving such signals could provide new insights in many studies. Moreover, the improved signal resolution in an overlapping spectrum would allow complex samples to be studied. Many times spectroscopic data is followed by model simulations which rely on the quality of the experimental signal. Denoising, by enhancing the accuracy of the data, can allow better comparisons with simulations.

### 3.5 Conclusion

We show that our NERD method can remove noise from experimental signals obtained from spectroscopy. It is able to accurately obtain signals at and above an *SNR* of 3, where standard signal processing methods can barely denoise data at *SNR*

of 30. Through this study, we show that signal denoising can be a powerful method to obtain high resolution spectra, thereby complementing the instrument sensitivity. This can allow studies to be conducted at low sample concentrations or studies that were not feasible due to long signal acquisition times. NERD reduces signal averaging times by at least two orders-of-magnitude. As NERD is not based on any prior information of the signal nor is tailored for any particular type of signal, it can be generally used in any spectroscopy.

## References

- (1) Pavia, D.; Lampman, G.; Kriz, G.; Vyvyan, J. *Introduction to Spectroscopy*; Cengage Learning, 2008.
- (2) Ernst, R. R. Sensitivity Enhancement in Magnetic Resonance. I. Analysis of the Method of Time Averaging. *Rev. Sci. Instrum.* **1965**, *36* (12), 1689.
- (3) Fodor, I. K. Denoising through Wavelet Shrinkage: An Empirical Study. *J. Electron. Imaging* **2003**, *12* (1), 151.
- (4) Wink, A. M.; Roerdink, J. B. J. M. Denoising Functional MR Images: A Comparison of Wavelet Denoising and Gaussian Smoothing. *IEEE Trans. Med. Imaging* **2004**, *23* (3), 374–387.
- (5) Cai, T.; Silverman, B. Incorporating Information on Neighbouring Coefficients into Wavelet Estimation. *Sankhyā Indian J. Stat. Ser. B* **2001**, *63* (2), 127–148.
- (6) Jun Jiang; Jian Guo; Weihua Fan; Qingwei Chen. An Improved Adaptive Wavelet Denoising Method Based on Neighboring Coefficients. In *2010 8th World Congress on Intelligent Control and Automation*; IEEE, 2010; pp 2894–2898.
- (7) Chen, G. Y.; Bui, T. D. Multiwavelets Denoising Using Neighboring Coefficients. *IEEE Signal Process. Lett.* **2003**, *10* (7), 211–214.
- (8) Chen, G. Y.; Bui, T. D.; Krzyzak, A. Image Denoising Using Neighbouring Wavelet Coefficients. *Integr. Comput. Aided. Eng.* **2005**, *12* (1), 99–107.
- (9) Dengwen, Z.; Wengang, C. Image Denoising with an Optimal Threshold and Neighbouring Window. *Pattern Recognit. Lett.* **2008**, *29* (11), 1694–1697.
- (10) Madhu, S.; Bhavani, H. B.; Sumathi, S.; Vidya, H. A. A Novel Algorithm for Denoising of Simulated Partial Discharge Signals Using Adaptive Wavelet Thresholding Methods. In *2015 2nd International Conference on Electronics and Communication Systems (ICECS)*; IEEE, 2015; pp 1596–1602.
- (11) Hussein, R.; Shaban, K. B.; El-Hag, A. H. Histogram-Based Thresholding in Discrete Wavelet Transform for Partial Discharge Signal Denoising. In *2015*

*International Conference on Communications, Signal Processing, and their Applications (ICCSPA'15)*; IEEE, 2015; pp 1–5.

- (12) Cetin, A. E.; Tofighi, M. Projection-Based Wavelet Denoising [Lecture Notes]. *IEEE Signal Process. Mag.* **2015**, 32 (5), 120–124.
- (13) Zhao, R.-M.; Cui, H. Improved Threshold Denoising Method Based on Wavelet Transform. In *2015 7th International Conference on Modelling, Identification and Control (ICMIC)*; IEEE, 2015; pp 1–4.
- (14) Ding, Y.; Selesnick, I. W. Artifact-Free Wavelet Denoising: Non-Convex Sparse Regularization, Convex Optimization. *IEEE Signal Process. Lett.* **2015**, 22 (9), 1364–1368.
- (15) Qiming Zhao; Wenzhan Dai. A Wavelet Denoising Method of New Adjustable Threshold. In *2015 IEEE 16th International Conference on Communication Technology (ICCT)*; IEEE, 2015; pp 684–688.
- (16) Madadi, Z.; Anand, G. V.; Premkumar, A. B. Signal Detection in Generalized Gaussian Noise by Nonlinear Wavelet Denoising. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2013**, 60 (11), 2973–2986.
- (17) Deng, G.; Liu, Z. A Wavelet Image Denoising Based on the New Threshold Function. In *2015 11th International Conference on Computational Intelligence and Security (CIS)*; IEEE, 2015; pp 158–161.
- (18) Bibina, V. C.; Viswasom, S. Adaptive Wavelet Thresholding & Joint Bilateral Filtering for Image Denoising. In *2012 Annual IEEE India Conference (INDICON)*; IEEE, 2012; pp 1100–1104.
- (19) Guo, S.; Wu, F.; Wei, W.; Guo, J.; Ji, Y.; Wang, Y. A Novel Wavelet Denoising Method Used for Droplet Volume Detection in the Microfluidic System. In *2013 IEEE International Conference on Mechatronics and Automation*; IEEE, 2013; pp 1732–1737.
- (20) Chen, G.; Xie, W.; Zhao, Y. Wavelet-Based Denoising: A Brief Review. In *2013 Fourth International Conference on Intelligent Control and Information Processing (ICICIP)*; IEEE, 2013; pp 570–574.
- (21) Donoho, D. L. De-Noising by Soft-Thresholding. *IEEE Trans. Inf. Theory*

- 1995**, *41* (3), 613–627.
- (22) Donoho, D. L.; Johnstone, J. M. Ideal Spatial Adaptation by Wavelet Shrinkage. *Biometrika* **1994**, *81* (3), 425–455.
- (23) Srivastava, M.; Anderson, C. L.; Freed, J. H. A New Wavelet Denoising Method for Selecting Decomposition Levels and Noise Thresholds. *IEEE Access* **2016**, *4*, 3862–3877.
- (24) Srivastava, M.; Georgieva, E. R.; Freed, J. H. A New Wavelet Denoising Method for Experimental Time-Domain Signals: Pulsed Dipolar Electron Spin Resonance. *J. Phys. Chem. A* **2017**, *121* (12), 2452–2465.
- (25) Bryant, J. H. Electron Paramagnetic Resonance Image Reconstruction for Arbitrary Projection Sampling and Regions of Interest, The University of Chicago, 2011.
- (26) Daubechies, I. Ten Lectures on Wavelets. *J. Acoust. Soc. Am.* **1993**, *93* (3), 1671.
- (27) Addison, P. S. *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance*; CRC Press: Philadelphia, 2002.
- (28) Donoho, D. L. De-Noising by Soft-Thresholding. *IEEE Trans. Inf. Theory* **1995**, *41* (3), 613–627.
- (29) Starck, J.-L.; Fadili, J.; Murtagh, F. The Undecimated Wavelet Decomposition and Its Reconstruction. *IEEE Trans. Image Process.* **2007**, *16* (2), 297–309.
- (30) Weil, J. E.; Bolton, J. R.; Wertz, J. E. *Electron Spin Resonance*; Wiley and Sons New York: Dordrecht, 1994.
- (31) Freed, J. H. New Technologies in Electron Spin Resonance. *Annu. Rev. Phys. Chem.* **2000**, *51* (1), 655–689.
- (32) Zhang, F.; Ma, L. MRI Denoising Using the Anisotropic Coupled Diffusion Equations. In *2010 3rd International Conference on Biomedical Engineering and Informatics*; IEEE, 2010; Vol. 1, pp 397–401.

- (33) Hofbauer, W.; Earle, K. A.; Dunnam, C. R.; Moscicki, J. K.; Freed, J. H. High-Power 95 GHz Pulsed Electron Spin Resonance Spectrometer. *Rev. Sci. Instrum.* **2004**, 75 (5), 1194–1208.
- (34) Freed, J. H. ESR and Molecular Dynamics. In *Biomedical EPR, Part B: Methodology, Instrumentation, and Dynamics*; Kluwer Academic Publishers-Plenum Publishers: New York; pp 239–268.

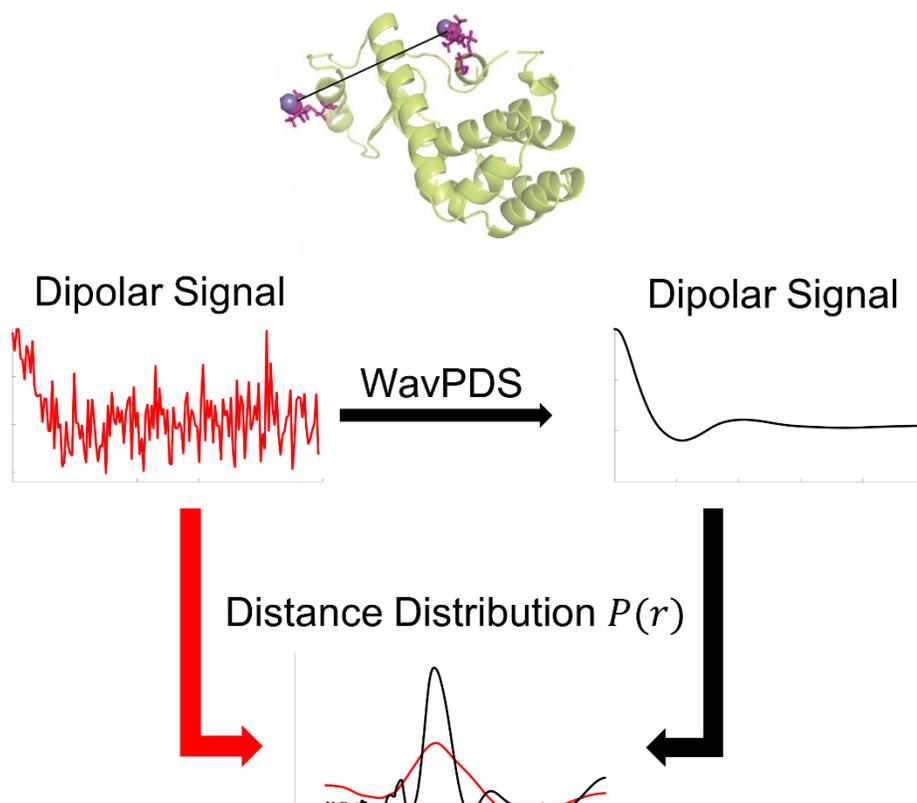
## CHAPTER 4

### A NEW WAVELET DENOISING METHOD FOR EXPERIMENTAL TIME-DOMAIN SIGNALS: PULSED DIPOLAR ELECTRON SPIN RESONANCE

#### Abstract

We adapt a new wavelet-transform based method of denoising experimental signals to pulse-dipolar electron-spin resonance spectroscopy (PDS). We show that signal averaging times of the time-domain signals can be reduced by as much as two orders of magnitude, while retaining the fidelity of the underlying signals when compared to noiseless reference signals. We have achieved excellent signal recovery when the initial noisy signal has an SNR  $\gtrsim 3$ . This approach is robust and is expected to be applicable to other time-domain spectroscopies. In PDS, these time-domain signals representing the dipolar interaction between two electron spin labels is converted into their distance distribution function  $P(r)$ , usually by regularization methods such as Tikhonov Regularization. The significant improvements achieved by using denoised signals for this regularization are described. We show that they yield  $P(r)$ 's with more accurate detail and yield clearer separations of respective distances, which is especially important when the  $P(r)$ 's are complex. Also, longer distance  $P(r)$ 's, requiring longer dipolar evolution times, become accessible after denoising. In comparison to standard wavelet denoising approaches, it is clearly shown that the new method (WavPDS) is superior.

## Graphical Abstract



\*This chapter was published in The Journal of Physical Chemistry A.

**Authors:** Madhur Srivastava, Elka R. Georgieva, and Jack H. Freed.

**Author Contributions:** Madhur Srivastava conducted research and wrote the paper, Elka R. Georgieva prepared the samples and performed the experiments, Jack Freed supervised the research and edited the paper.

## 4.1 Introduction

### 4.1.1 Background

Pulsed dipolar ESR spectroscopy (PDS) is a powerful biophysical method for the study of the structure and function of biological systems.<sup>1-9</sup> It has been applied mainly to proteins and protein complexes, which either possess endogenous paramagnetic centers<sup>10-12</sup> or engineered sites to which paramagnetic tags (spin-labels) are covalently attached.<sup>13,14</sup> PDS measures the strength of the magnetic dipole-dipole interaction between electron spins,<sup>1,2,4</sup> usually carried by the attached spin-labels. However, in many cases studying the system of interest by PDS is a challenging task due to several factors, such as low protein concentrations available in a particular experiment, and/or short phase memory relaxation times ( $T_m'$ ) causing a fast decay of the dipolar signal over a few microseconds, particularly in the case of membrane proteins. All these obstacles lead to dipolar evolution signals in the time-domain often with insufficient signal-to-noise ratios (SNR). Thus, very long time-averaging of the PDS signals is frequently needed in order to obtain reliable data for further analysis to reconstruct accurate distances and distance distributions that report on the protein structure. Several approaches have been used to overcome these challenges: solvent deuteration,<sup>3,15,16</sup> and also partial<sup>17</sup> or complete<sup>18</sup> protein deuteration; the building of spectrometers with increasing sensitivity, either by several research groups<sup>19-22</sup> or commercially; and the development of new pulse sequences that enhance the accessible distance range and allow samples with reduced protein concentration to be

used.<sup>23,24</sup> Yet, new methods for improving the SNR of PDS experiments remain needed.

Herein we report our development of a novel wavelet transform denoising method,<sup>25</sup> which provides significant noise removal in time-domain signals while preserving the fidelity of the underlying original signals. It is a powerful approach and can be used to obtain good PDS data from noisy signals. The method provides multiresolution analysis that enables time-frequency decomposition of the signal, i.e., it informs which frequencies occur at what time instances, which helps to separate the signal from the noise.<sup>26-30</sup> The method significantly improves on existing wavelet denoising methods by providing a much larger increase in SNR with virtually no distortion of the underlying signal.

While we focus on the application to time-domain ESR, we note that it is equally applicable to other physical chemical techniques that examine time-domain signals.

To demonstrate the new denoising method as applied to PDS, we used data for doubly spin-labeled cysteine mutants of T4 Lysozyme (T4L) as well as model data of unimodal and bimodal Gaussian distance distributions.

We show that using this new method of denoising the time-domain signal improves the stability of the reconstruction of the inter-spin distance distributions, and therefore the reliability of the results. Also, we demonstrate that signal acquisition times can be

reduced by at least an order of magnitude and still yield high quality results after denoising. The results from denoising of the bimodal model and bimodal experimental data show that the method can be applied with very high accuracy to more complex PDS signals typical of multiple inter-spin distance distributions. Thus, the challenge of obtaining PDS time-domain signals with high SNR is much reduced, so that systems at low spin-concentration and systems that produce rapidly decaying PDS signals due to short  $T_m$ 's become more amenable to study. We strongly believe that the denoising method described here will greatly benefit the field of PDS studies. In addition, this method should apply to other experimental fields where noisy time-decay signals are observed, such as Nuclear Magnetic Resonance Spectroscopy,<sup>31</sup> Infrared Spectroscopy,<sup>32</sup> Optical Spectroscopy,<sup>33</sup> and others.

#### 4.1.2 Conversion of Pulse Dipolar Signals to Distance Distributions

The signal from an ensemble of proteins with two spin labels each is given as

$$\int_{R_{min}}^{R_{max}} \kappa(r, t) P(r) dr = S(t) \quad (4.1)$$

where  $\kappa(r, t)$  represents an orientationally averaged signal from a spin-pair at a given  $r$ ,  $P(r)$  is the distribution in distance between the spin pairs, and  $S(t)$  is the PDS signal. In discrete matrix form, it can be written as

$$KP = S \quad (4.2)$$

which is especially appropriate for discrete time data acquisition. The signal  $S(t)$  is acquired from the PDS experiment. To obtain the distance distribution  $P(r)$ , the inversion of  $K$  is required (i.e.,  $P = K^{-1}S$ ), which is an ill-posed problem. As the determinant of  $K$  is zero or near zero, there are many possible solutions for the distance distribution  $P$ . Also this inversion is easily corrupted by the noise in the experimental results. To overcome these problems and obtain a stable and desirable solution  $P$ , methods like Pake pattern function,<sup>34</sup> Direct Conversion,<sup>35</sup> Singular Value Decomposition (SVD),<sup>36</sup> Model Fitting,<sup>37-48</sup> Polynomial Interpolation,<sup>49</sup> Coordinate Transformation,<sup>50</sup> Tikhonov Regularization (TIKR),<sup>36,43,49,51,52</sup> and Regularization using Monte Carlo calculations<sup>53</sup> have been used. The basic SVD used to invert  $K$  and to obtain  $P$  is given by

$$P = \sum_{i=1}^p \frac{u_i^T S}{\sigma_i} v_i \quad (4.3)$$

where  $u_i$  and  $v_i$  are column vectors of unitary matrices  $U$  and  $V$ , obtained in the SVD of the matrix  $K$ , and the  $\sigma_i$  are the individual singular values of the diagonal matrix  $\Sigma$  obtained from  $K$ , and  $p$  is the length of the input signal  $S(t)$ .<sup>36</sup> Given the instability of equation 4.3, the TIKR method is often used.<sup>36,43,49,51</sup> It minimizes the function

$$\phi_{TIKR}[P] \equiv \min \|KP - S\|^2 + \lambda^2 \|LP\|^2 \quad (4.4)$$

where  $\lambda$  is the regularization parameter and  $L$  is the differentiation operator. Equation 4.4 leads to the unique distance distribution given by

$$P_\lambda = (K^T K + \lambda^2)^{-1} K^T S \quad (4.5)$$

For practical implementation, equation 4.5 can be rewritten in terms of the SVD (cf. Eq. 4.3) as

$$P_\lambda = (\Sigma^2 + \lambda^2 I)^{-1} V \Sigma U^T S \quad (4.6)$$

Equation 4.6 can be rewritten and solved as a sum of singular value contributions:<sup>36,51</sup>

$$P_\lambda = \sum_{i=1}^{\text{Rank}(K)} f_i \frac{u_i^T S}{\sigma_i} v_i \quad (4.7)$$

where  $f_i \equiv \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2}$

Fig. 4.1A shows the block diagram of the standard approach.

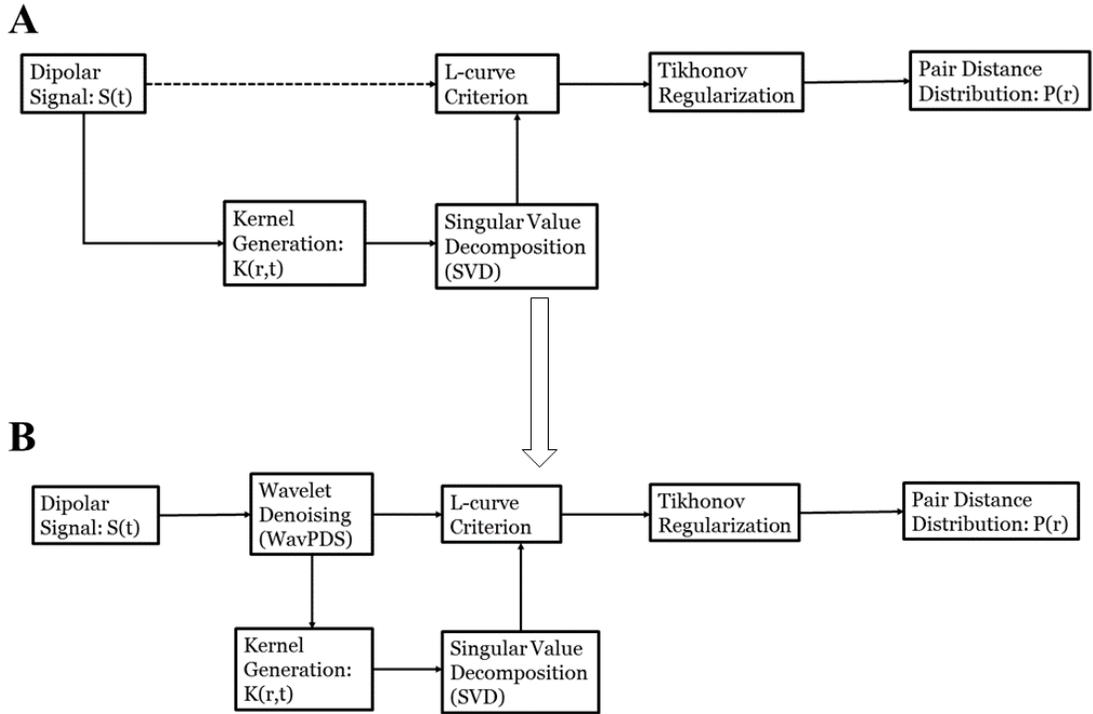
Limitations:

- 1) The ill-posed problem of Eq. 4.3 is highly sensitive to noise present in the signal, for which the method of Eq. 4.7 has some effectiveness. In suppressing noise, this method tends to lead to broader distance distributions. The broadening of the distance distribution is attributed to the removal of the

contributions from the smaller singular values, for which  $\sigma_i^2 \ll \lambda^2$ , by the filter function  $f_i$ . But incorporating contributions of smaller singular values, which are substantially corrupted by noise, would yield an unstable solution. Hence, there is a trade-off between an unstable solution and a reduction of resolution in the distance distribution due to the presence of noise, which requires the removal of contributions of smaller singular values.

- 2) It is difficult to obtain a high SNR using longer dipolar evolution times. At longer evolution times, the SNR becomes poorer increasing the instability in the distance distribution. Hence, smaller evolution times are typically used with relatively higher SNR to obtain a more stable distribution, but at a loss of resolution in the distance distribution. This is especially a problem for longer distances.

In this paper, we present our method to denoise the dipolar signal in order to retrieve a noise-free signal, and subsequently yield stable and accurate distance distributions. Fig. 4.1 shows the block diagram of the TIKR process with and without applying wavelet denoising.



**Figure 4.1:** Block diagrams for determining distance distribution  $P(r)$  from the dipolar signal by Tikhonov regularization using L-curve criterion. **A)** Standard approach; **B)** New approach after WavPDS denoising.

### 4.1.3 Wavelet Transform and Discrete Wavelet Transform

A wavelet transform (WT) provides the time-frequency decomposition of a signal. This helps to separate noise and signal, as they typically do not possess the same time-frequency patterns. The wavelet transform of a signal  $S(t)$  is given as<sup>26,28,29</sup>

$$F(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{+\infty} S(t) \psi^* \left( \frac{t - \tau}{s} \right) dt \quad (4.8)$$

where  $\tau$  (a time) and  $s$  (an inverse frequency) are respectively known as the translation and scale parameters,  $S(t)$  is the input signal,  $F(\tau, s)$  is the WT of the signal as a function of  $\tau$  and  $s$ , and  $\psi^*(t)$  is the complex conjugate of the wavelet function  $\psi(t)$ ; for real wavelets,  $\psi^*(t) = \psi(t)$ . As can be seen from equation 4.8,  $\tau$  and  $s$  allow the study of time-frequency information of the input signal by varying the window width  $s$  and its translation  $\tau$ . However, it is computationally cumbersome to perform the WT calculation for all the possible values of  $\tau$  and  $s$ .

By analogy to the Fast Fourier Transform, the Discrete Wavelet Transform (DWT) is used to calculate the wavelet transform in a computationally less expensive manner. The DWT is given as<sup>26,29,54</sup>

$$D_j[n] = \sum_{m=0}^{p-1} S[t_m] 2^{j/2} \psi[2^j t_m - n] \quad (4.9)$$

where now  $s = 2^{-j}$  and  $\tau = n2^j$ , with  $j$  and  $n$  integers,  $t_m$  is the discretized time,  $S[t_m]$  is the discretized input signal,  $p = \text{length}(S[t_m])$ , and  $D_j[n]$  is the discrete wavelet transform (also called the Detail component) of  $S[t_m]$  at scale  $2^j$ , also known as the decomposition level  $j$ . This represents an expansion of the function  $S(t)$  in an orthonormal set of wavelet functions expressed in Equation 4.10.<sup>29</sup>

$$\sum_{m=1}^L \psi\left(\frac{t_m - n2^j}{2^j}\right) \psi\left(\frac{t_m - n'2^{j'}}{2^{j'}}\right) = \begin{cases} 1, & \text{if } j = j' \text{ and } n = n' \\ 0, & \text{otherwise} \end{cases} \quad (4.10)$$

Here  $L$  is the discrete length of the wavelet function that is a characteristic of the particular wavelet used. The number of decomposition levels allowed is  $N$ , where  $N = \lceil \log_2 p \rceil$ , i.e.,  $1 \leq j \leq N$ , where  $N$  refers to the lowest frequency sub-band. The dyadic scale (i.e.,  $s = 2^{-j}$ ) is selected so that the Detail components  $D_j[n]$  represent non-overlapping frequency bands.<sup>30</sup> However, at decomposition levels 1 to  $j$ , only frequency sub-bands represented by them are wavelet-transformed, requiring that the remaining frequency band to reconstruct the original signal also be obtained. Even at  $j = N$ , there is a low-frequency sub-band that is not covered because  $j$  is still constrained by the discrete finite signal of step length given by  $p$ . The remaining low frequency information contained in the wavelet transform of the signal is given by the Approximation component, which is defined as

$$A_j[n] = \sum_{m=0}^{p-1} S[t_m] 2^{j/2} \phi[2^j t_m - n]$$

(4.11)

where  $A_j[n]$  is the Approximation component at the  $j^{th}$  decomposition level, and  $\phi[t_m]$  is the discrete scaling function that allows the calculation of the low frequency part of the wavelet transform of the input signal. The wavelet function  $\psi[2^j t_m - n]$  and the scaling function  $\phi[2^j t_m - n]$  derived from it are orthogonal and of the same length.<sup>26,29</sup> The input signal can be recovered using the inverse discrete wavelet transform (IDWT) from the  $D_j[n]$  of Eq. 4.9 and the  $A_j[n]$  of Eq. 4.11.<sup>26,29,54</sup>

We show in Fig. 4.S1 the decomposition of an actual PDS signal into a sequence of Detail and Approximation components.

## 4.2 Method

### 4.2.1 Wavelet Denoising

The Detail components in the DWT allow the separation of signal and noise.<sup>54,55</sup> Moreover, given a good choice of wavelet family, the signal, which is coherent, will have a large magnitude in a few wavelet coefficients and will occur at the same location at the different Detail components. By contrast, random noise will have many wavelet coefficients with small magnitudes that vary in location at the different Detail components. To denoise the signal in the DWT, the following four general steps are required,<sup>56</sup> which we described in terms of our new method.<sup>25</sup>

- a) The selection of an appropriate wavelet. We use “db6” from the Daubechises wavelet family<sup>57</sup> in this study. They are illustrated in Fig. 4.S2. This wavelet correlates very well with the signal properties of decaying oscillatory time-domain signals. This results in effective separation of signal and noise by increasing the difference between their wavelet coefficient patterns.
- b) The selection of the maximum decomposition level needed to properly denoise the Detail components and the Approximation component. Smaller or greater maximum decomposition levels can result in under-denoising or signal distortion, respectively. We use the criterion  $\{S_j < T_r, S_{j+1} > T_r, \text{ when } k = j\}$ , where  $k$  is the maximum decomposition level needed,  $T_r$  is the criterion to distinguish between noisy and noise-free Detail components discussed below, and  $S_j$  is the “peak-to-sum ratio”, given as

$$S_j = \frac{\max(D_j)}{\sum_{n=1}^{q_i} D_j[n]} \quad (4.12)$$

Thus,  $S_j$  is a measure of the sparsity of the  $j^{th}$  decomposition level. It is calculated separately for negative and positive coefficients ( $S_{j,L}$  and  $S_{j,H}$ , respectively) at the  $j^{th}$  decomposition level.

- c) To separate the noise coefficients from the signal coefficients, noise thresholds for each Detail component are selected. As mentioned above, the noise coefficients will have smaller magnitude than the signal coefficients, i.e., the maximum magnitude of noise coefficients should be less than the minimum magnitude of signal coefficients. We use two thresholds, lower and upper, for negative and positive coefficients. They are calculated as  $\lambda_{j,M} = \mu_j - \kappa_{j,M}\sigma_j$  where  $M = L$  or  $H$  corresponding to the lower and upper thresholds, respectively, for the Detail component at the  $j^{th}$  Decomposition level,  $\mu_j$  and  $\sigma_j$  are the mean and standard deviation, respectively of the Detail component at the  $j^{th}$  Decomposition level, and  $\kappa_{j,L}$  and  $\kappa_{j,H}$  are adjustable parameters; our previous paper<sup>25</sup> describes how to obtain these values based upon the associated  $S_{j,L}$  and  $S_{j,H}$  values.
- d) The selection of the noise thresholding function to apply the noise thresholds: We have devised a condition in ref. 25 that leads to the use of wavelets such that a coefficient in a Detail component is either a noise or a signal coefficient provided the SNRs are above some minimum value. We have carefully

confirmed that the wavelets in this study as well as those in ref. 25 satisfy this condition, resulting in successfully denoising at a minimum SNR compared to other standard wavelet families. This enables us to use as the thresholding function:

$$D'_j[n] = \begin{cases} 0, & \text{for } \lambda_{j,L} < D_j[n] < \lambda_{j,H} \\ D_j[n], & \text{otherwise} \end{cases} \quad (4.13)$$

where  $D_j[n]$  and  $D'_j[n]$  are the noisy and denoised Detail component, respectively, at the  $j^{\text{th}}$  Decomposition level.

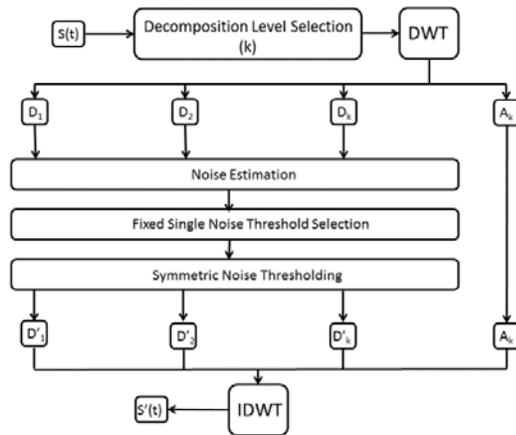
In addition, we carry out two more steps tailored to the needs of a rapidly decaying signal versus time. First, the signal is flipped from left to right in time (i.e., time is reversed) before taking the DWT to avoid distorting the initial ( $t = 0$ ) signal. Reversing the signal results instead in having small, near-zero, magnitudes at the start, allowing more accurate DWT of the stronger part of the signal. The DWT of the flipped signal, of course, contains the same information as that of the non-flipped signal.

Second, the signal is divided into two parts: higher SNR and lower SNR regions. Although the noise content is the same, the higher SNR part for earlier times contains signal that is less affected by noise, while noise tends to be dominant in the low SNR part. Therefore, the number of Detail components to denoise is different for both.

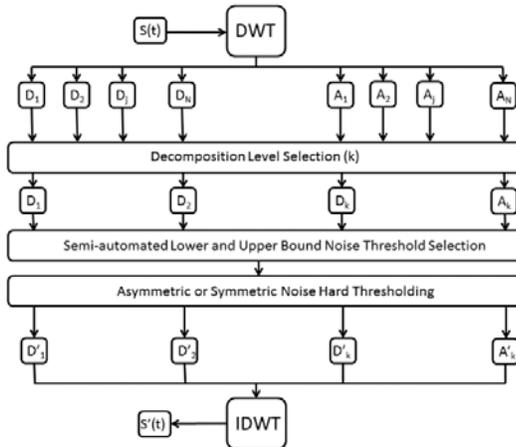
Moreover, the separation into lower and higher SNR parts avoids an overlap between the signal wavelet coefficients of the higher SNR part and the noise wavelet coefficients of the lower SNR part, preventing under-denoising or signal distortion. Fig. 4.2 compares in block diagram form a standard wavelet denoising method (2A) with our new denoising process (2B)<sup>25</sup> and the modified version for time-domain signals (2C).

We call our denoising method with the additional modifications “WavPDS” for PDS signals.

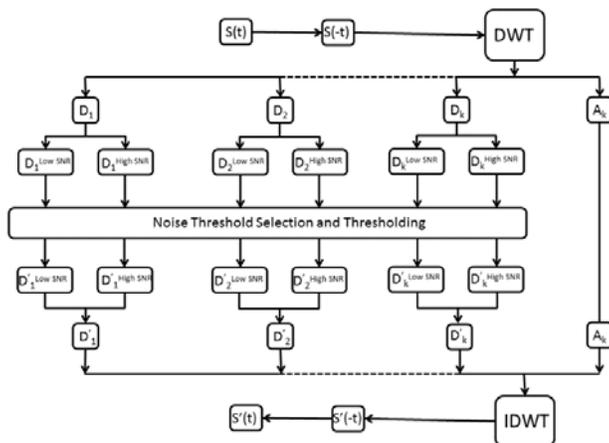
## A Standard Wavelet Denoising Method



## B New Method



## C WavPDS Method



**Figure 4.2:** **A)** The block diagram of a standard wavelet denoising method; **B)** The block diagram of the new wavelet denoising method;<sup>25</sup> **C)** The block diagram of WavPDS.  $S(t)$  and  $S'(t)$  are the noisy and denoised signal, respectively;  $D_j$  and  $D'_j$  are the noisy and denoised Detail components at the  $j^{th}$  Decomposition level, respectively;  $A_j$  and  $A'_j$  are the noisy and denoised Approximation components at the  $j^{th}$  Decomposition level, respectively; DWT and IDWT represents Discrete Wavelet Transform and Inverse Wavelet Transform, respectively. (Figs. 4.2A and 4.2B reprinted with permission from ref 25. © 2016 IEEE.)

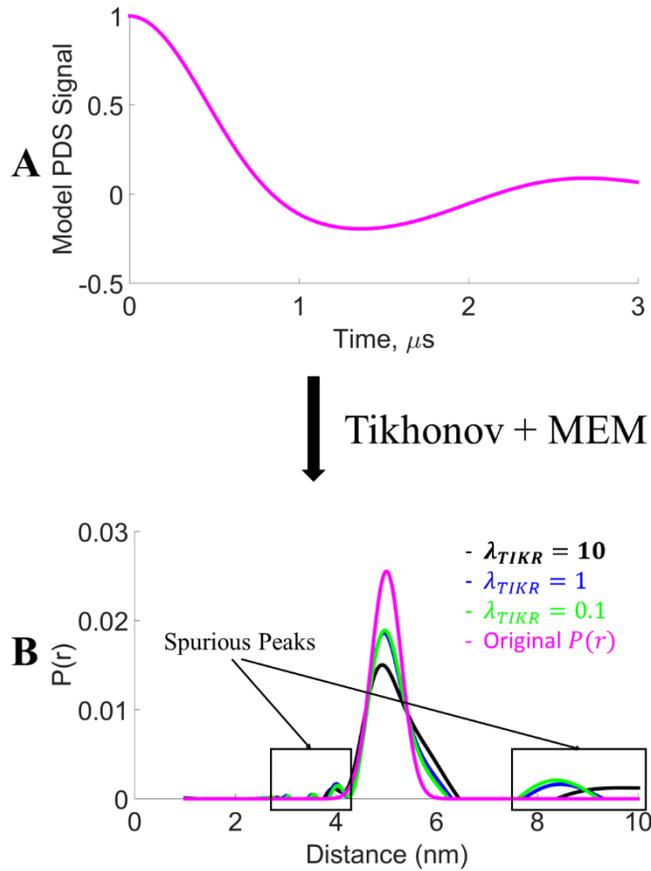
## 4.2.2 Difference between the New Method and Standard Methods

Standard wavelet denoising methods are vulnerable to signal distortion and under-denoising, as they are not able to effectively remove some noise coefficients and/or may remove some signal coefficients in the wavelet domain. (cf. Figs. 4.S3-4.S8).<sup>27,56,58,59</sup> WavPDS avoids this by using our new method<sup>25</sup> which objectively selects the number of decomposition levels to denoise, calculates the adaptive noise thresholds based on statistics of wavelet coefficients using a new formula, uses separate noise thresholds for positive and negative wavelet coefficients, and removes low frequency noise by denoising the approximation component (cf. the block diagrams in Figs. 4.2A, 4.2B & 4.2C).<sup>25</sup> The upshot of these features as discussed in Section 4, Item 5 is that WavPDS is more successful in removing the noise coefficients in the Detail components representing high-frequency sub-bands and only WavPDS removes the lowest frequency noise present in the final Approximation component.

## 4.2.3 Spurious Peaks

Although noise in the signal is a source of spurious or unwanted peaks in the distance distribution, the removal of the smaller of singular values, which are corrupted by the noise, also contributes to the spurious peaks. To demonstrate this, we generated a model unimodal Gaussian distribution centered at 5 nm with a standard deviation of 0.3 nm. Then, a model signal was generated for 3  $\mu$ s using this model distance distribution. The distance distributions were reconstructed by using different  $\lambda_{TIKR}$

values in the TIKR<sup>36</sup> followed by the Maximum Entropy Method (MEM).<sup>51</sup> MEM was used largely to remove negative values of  $P(r)$  that emerged from the TIKR.<sup>51</sup> Fig. 4.3 shows different  $P(r)$ 's generated using different  $\lambda_{TIKR}$  values. The original  $P(r)$  (in pink) is given by the model and is also obtained from the model  $S(t)$  as expected (cf. Fig. 4.3A) after TIKR using the L-curve<sup>36</sup> (cf. Sections 3.1 and 3.2). Fig. 4.3 shows that the distribution narrows with smaller  $\lambda_{TIKR}$  values, i.e., more contributions from smaller singular values. This result is also expected. However, Fig. 4.3 also shows the presence of spurious peaks despite using a noise-free model. These spurious peaks (and negative values not shown in Fig. 4.3 because MEM has been applied) are due to the removal of contributions of the smaller singular values. Thus, spurious peaks are generated not only due to the presence of noise, but also to the absence of contributions of the smaller singular values, which are expunged by the TIKR method for singular values where for  $\sigma_i^2 \ll \lambda_{TIKR}^2$ .



**Figure 4.3:** Comparison of distance distributions with different  $\lambda$  values for Tikhonov regularization of the model signal. The model signal was generated from a Gaussian distribution centered at 5 nm with standard deviation of 0.3 nm. (Note: MEM<sup>51</sup> was used to constrain  $P(r) \geq 0$ ).

#### 4.2.4 Objective Measures

The SNR,  $\chi^2$ , and Structural Similarity Index Measure (SSIM)<sup>60,61</sup> are used as objective measures to compare the results of noisy and denoised signals with respect to a reference signal. The SNR provides the estimate about the noise in the noisy or denoised signal,  $\chi^2$  provides the closeness of fit, and the SSIM provides the similarity of the noisy or denoised signal with respect to a reference signal in terms of their shape and structure. The SNR is calculated as

$$SNR = \frac{Signal_{Peak}}{Noise_{RMS}} \quad (4.14)$$

Also  $\chi^2$  is the root mean square (RMS) error defined as

$$\chi^2 \equiv \sqrt{\frac{1}{p} \sum_{i=1}^p (Y[i] - X[i])^2} \quad (4.15)$$

where  $X$  and  $Y$  are input and reference signals, respectively. The SSIM is calculated as

$$SSIM(X, Y) = \frac{(2\mu_X\mu_Y + c_1)(2\sigma_{XY} + c_2)}{(\mu_X^2 + \mu_Y^2 + c_1)(\sigma_X^2 + \sigma_Y^2 + c_2)} \quad (4.16)$$

where  $\mu$  and  $\sigma$  are mean and standard deviation, respectively, and  $c_1$  and  $c_2$  are positive integers used to stabilize each term. They are given in the standard MATLAB

package. The SSIM values range from -1 to 1, where an SSIM of 1 means  $X$  is identical to  $Y$ .<sup>60</sup>

#### 4.2.5 ESR Method and Sample Preparation

DEER experiments at 17.3 GHz were conducted on several samples of double cysteine mutants of bacteriophage T4 lysozyme (T4L) spin-labeled with MTSL ((1-Oxyl-2,2,5,5-tetramethyl-D3-pyrroline-3-methyl) Methanethiosulfonate). These T4L mutants were generated by site-directed mutagenesis using the Quick-Exchange Multi Site-Directed Mutagenesis Kit (Agilent Technologies, Stratagene Products Division), expressed in *E. coli* BL21(DE3) cells, and purified and spin-labeled as previously described.<sup>62</sup> After spin-labeling and removal of unreacted MTSL, the protein concentration was determined from the UV absorbance at 280 nm.<sup>62</sup> The following samples of spin-labeled T4L were studied: (i) SAMPLE 1—63  $\mu\text{M}$  of mutant 44C/135C; (ii) SAMPLE 2—mixture of mutants 8C/44C and 44C/135C at concentrations of 44  $\mu\text{M}$  and 47  $\mu\text{M}$ , respectively. In both cases, the buffer composition was—25 mM Tris/25 mM MOPS pH 7.6, 1 mM EDTA and 10% Gly (w/v); before freezing for PDS measurements, an extra 20% Gly-d8 (w/v) was added to the samples. Therefore, it should be noted that the protein concentrations given above are those before the extra 20% Gly-d8.

Four-pulse DEER<sup>2,63</sup> measurements at 17.3 GHz and 60 K were performed using a home-built Ku-band pulse spectrometer under standard experimental conditions.<sup>1,3,19</sup>

The spectrometer tune-up was, as usual, done by using both the primary and refocused echoes, which are much stronger than the DEER signal, so there was no difficulty tuning up even for a very dilute sub- $\mu\text{M}$  concentration sample of labeled protein (cf. Fig. 4.S9). The  $\pi/2$ - $\pi$ - $\pi$  DEER pulse widths were 16 ns, 32 ns and 32 ns, respectively, and the  $\pi$  pump pulse was 32 ns. A frequency separation of 70 MHz between detection and pump pulses was used. The detection pulses were applied at the low-field edge and the pump pulse was positioned near the center of nitroxide spin-label spectrum, close to the maximum. DEER time-domain signals of 4-5  $\mu\text{s}$  dipolar evolution times were collected. Three data-sets at different DEER signal averaging times were produced for each sample, resulting in widely different SNR. From these time-domain DEER data, inter-spin distance distributions were reconstructed using TIKR<sup>36</sup> and refined by MEM<sup>51</sup> to constrain  $P(r) \geq 0$ . Since SAMPLE 1 contained just one doubly spin-labeled protein mutant, it produced a unimodal distance distribution, whereas SAMPLE 2 produced a bimodal distance distribution due to the superposition of two unimodal distributions.

## 4.3 Results

### 4.3.1 Model Data

#### 4.3.1.1 Unimodal Distance Distribution

A 3  $\mu\text{s}$  signal was generated for a unimodal Gaussian distribution centered at 5 nm with a standard deviation of 0.3 nm. Gaussian noise was added to the model signal to generate SNRs of 3, 10, and 30. Denoising was performed on the noisy data, resulting

in SNRs of the denoised signals of 278, 848, and 1883, respectively. Fig. 4.4A shows the denoised time-domain signals along with the original noisy signals, while Fig. 4.4B compares the denoised signal with the model. All curves are normalized. It can be seen that for SNR 10 and 30, the denoising is able to (nearly) perfectly recover the model data. At SNR 3, the denoised signal shows only a small discrepancy with respect to the model. Table 4.1 shows the  $\chi^2$ , SNR, and SSIM values of the noisy signals and how they are greatly improved by denoising. The SNR is improved by almost two orders of magnitude and the  $\chi^2$  values are also greatly reduced. The SSIM values are significantly increased, approaching unity from much poorer comparisons of the noisy signals to the model.

The distance distributions of the model, noisy and denoised signals were obtained using TIKR<sup>36</sup> followed by MEM.<sup>51</sup> The regularization parameter value  $\lambda_{TIKR}$  was first obtained using the L-curve method.<sup>36,64–66</sup> However, at low SNR (i.e., the noisy signals), the L-curve method is not very reliable<sup>36</sup> and led either to a very broad distribution or to an unstable solution. Moreover, it was observed that the  $\lambda_{TIKR}$  value obtained from the L-curve using the criterion of maximum curvature was not at the elbow of the L-curve for these noisy signals, since there was no well-defined elbow. In fact, for SNRs of 3 and 10 the L-curve plot was very distorted. Therefore, the  $\lambda_{TIKR}$  value was then manually selected for the noisy signals (after trial and error) for optimal distance distribution i.e.,  $\lambda_{TIKR}^{OPT}$ . The optimal distance distribution was easy to determine as it could be compared with the model distribution. A larger  $\lambda_{TIKR}$  yielded a broader distance distribution, whereas a smaller  $\lambda_{TIKR}$  yielded an unstable

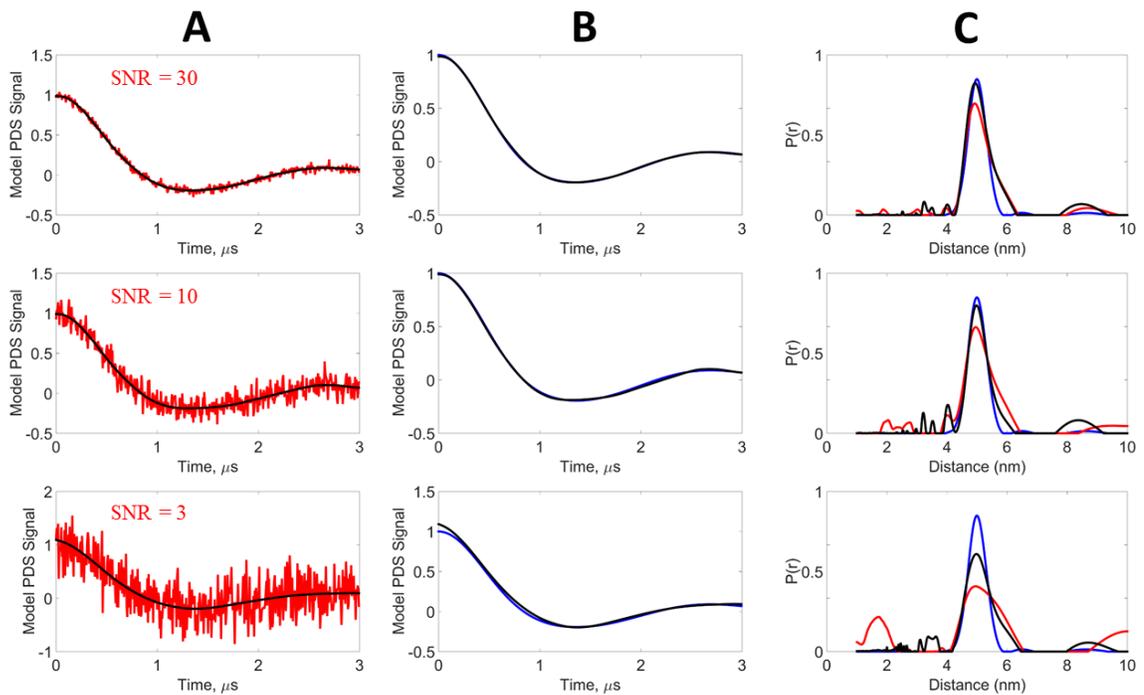
distribution and/or enhanced artificial peaks. However, once denoised, the standard shape of the L-curve was restored (i.e., a clear elbow), so the optimal  $\lambda_{TIKR}$  selected for the denoised signal was obtained directly from the L-curve. For the noisy signals, the  $\lambda_{TIKR}^{OPT}$  values obtained manually decreased after denoising ( $\lambda_{TIKR}^{OPT} = 15 \rightarrow 0.01$ ,  $\lambda_{TIKR}^{OPT} = 10 \rightarrow 0.5$ , and  $\lambda_{TIKR}^{OPT} = 5 \rightarrow 0.5$  for SNR of 3, 10, and 30, respectively). For the original model signal, the optimal  $\lambda_{TIKR}^{OPT} = 1 \times 10^{-6}$  was obtained using the L-curve method. Since the model data are noise-free,  $\lambda_{TIKR}$  is near zero in order to eliminate the contributions of the zero singular values that arise because the  $K$  matrix is singular.<sup>36</sup> While they should be exactly zero, due to computational error they are only near zero. The denoised signals consistently have much smaller optimal  $\lambda_{TIKR}$  compared to the noisy signals. This means that, when denoised, a significant number of contributions from the smaller singular values could be included according to Eq. 4.7.

Fig. 4.4C compares the distance distributions from the noisy, denoised, and model signals. At SNR = 30, the distance distribution of the denoised signal and the model are very nearly identical, whereas that for the noisy signal has a slightly reduced normalized intensity and is somewhat broader at its larger distances. For SNRs of 3 and 10, the distance distributions  $P(r)$  of the denoised signals accurately provide the peak location and distribution width, whereas the  $P(r)$  of the noisy signals possess broader peaks of lower amplitude. Especially for an SNR of 3, the distance distribution is no longer reliable.



**Table 4.1:** Comparison of Noisy and WavPDS Denoised signals with respect to Model signal - Unimodal Distribution at SNRs of 30, 10, and 3 (cf. Fig. 4.4).

	<b>SNR</b>	$\chi^2$	<b>SSIM</b>
<b>Noisy → Denoised</b>	30 → 1880	0.03 → 0.0005	0.5836 → 0.994
	10 → 850	0.10 → 0.0011	0.1194 → 0.959
	3 → 276	0.33 → 0.0036	0.010 → 0.938



**Figure 4.4:** Model Data - Unimodal Distribution: **Blue** – Model Signal (Reference), **Red** – Noisy Signal, **Black** – WavPDS Denoised Signal. The model signal was generated from a Gaussian distribution centered at 5 nm with a standard deviation of 0.3 nm. White Gaussian noise was added to generate the Noisy signals at SNRs 30, 10 and 3. **A)** Comparison of Noisy and Denoised signals; **B)** Comparison of the reference signal with the Denoised signal obtained by WavPDS; **C)** Distance Distributions from Noisy, Denoised, and Reference signal.

### 4.3.1.2 Bimodal Distance Distribution

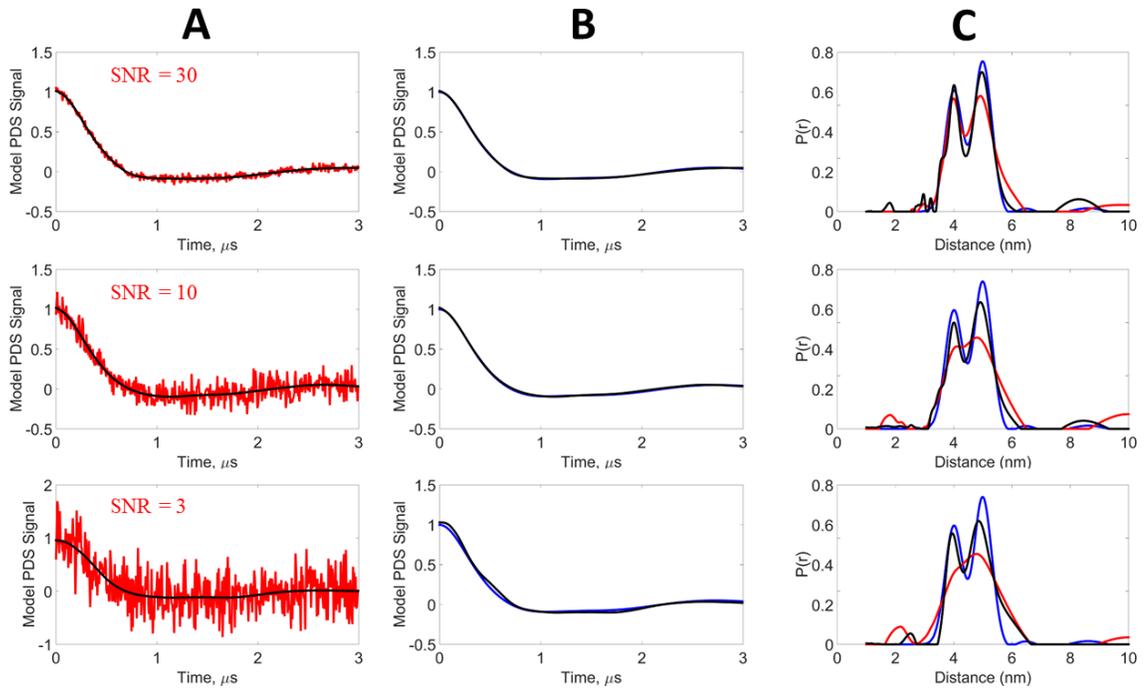
A 3  $\mu\text{s}$  signal was generated for a bimodal Gaussian distribution, centered at 4 nm and 5 nm, each with a standard deviation of 0.3 nm. The peak height at 4 nm was 80% of the peak height at 5 nm. Similar to the unimodal model, Gaussian noise was added to the model signal to generate SNRs of 3, 10, and 30, then was denoised. Table 4.2 shows the  $\chi^2$ , SNR, and SSIM values of the noisy and denoised signals. Figs. 4.5A and 4.5B show the denoised signals along with noisy and model signals, all of which are normalized. For the SNR of 30, the denoising is able to virtually perfectly recover the model data; the SSIM is very close to unity. For the SNRs of 10 and 3, the denoising is almost as good, as reflected in the SSIM values as well as in the  $\chi^2$ 's. The denoised signals significantly increase in SNR by about two orders of magnitude for all the noisy signals, just as we found for the unimodal case.

Similar to example 3.A.1, the L-curve plot  $\lambda_{TIKR}$  value needed to be manually adjusted for noisy signals, whereas the optimal  $\lambda_{TIKR}$  for the denoised signals was obtained directly from the L-curves (cf. Fig. 4.S10). The optimal  $\lambda_{TIKR}$  values were reduced from those for the noisy signals to those for the denoised signals ( $\lambda_{TIKR}^{OPT} = 60 \rightarrow 10$ ,  $\lambda_{TIKR}^{OPT} = 20 \rightarrow 3$ , and  $\lambda_{TIKR}^{OPT} = 5 \rightarrow 1$  for SNR 3, 10, and 30, respectively; cf. Fig. 4.S10). For the model signal itself, the optimal  $\lambda_{TIKR} = 5 \times 10^{-6}$  was simply obtained using the L-curve.

In Fig. 4.5C, the distance distributions of noisy, denoised, and model signals at SNRs 3, 10, and 30 are shown. At SNR = 30, the  $P(r)$  of the denoised signal is very close to the model  $P(r)$ , whereas the  $P(r)$  of the noisy signal has a broader distribution and lower peak height. Moreover, the first peak incorrectly appears to be about the same height as the second peak. For an SNR of 10, the noisy signal results in a broad distribution with the two peaks almost indistinguishable. Meanwhile, the denoised signal is able to resolve both peaks along with maintaining their height difference. The noisy signal at SNR = 3 provides only a single broad distribution, thereby failing to resolve the peaks. On the other hand, its denoised version is successfully able to clearly resolve the two peaks.

**Table 4.2:** Comparison of Noisy and WavPDS Denoised signals with respect to Model signal - Bimodal Distribution at SNRs 30, 10, and 3 (cf. Fig. 4.5).

	<b>SNR</b>	$\chi^2$	<b>SSIM</b>
<b>Noisy → Denoised</b>	30 → 3186	0.03 → 0.0003	0.5412 → 0.997
	10 → 1850	0.10 → 0.0005	0.1068 → 0.953
	3 → 378	0.33 → 0.0026	0.0077 → 0.945

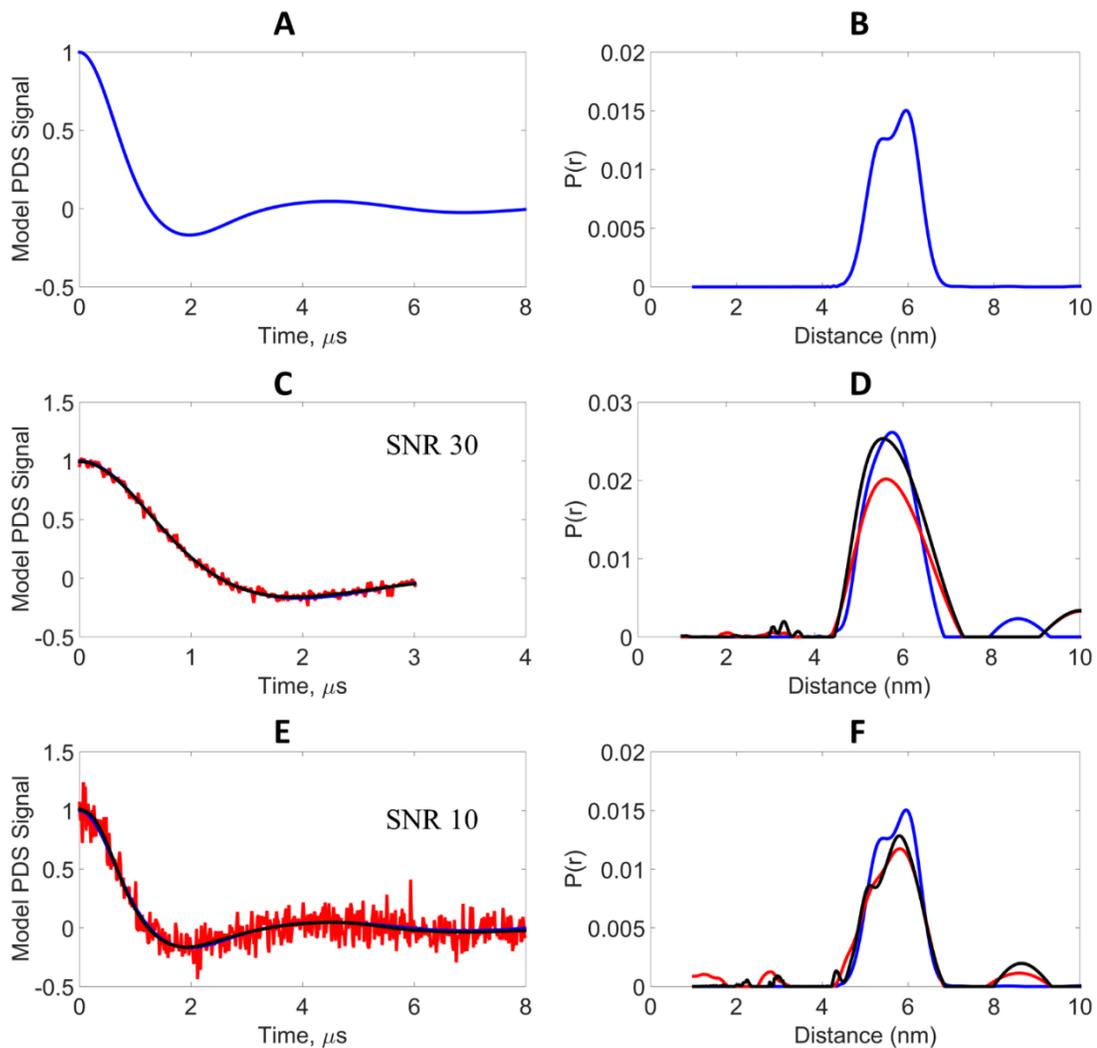


**Figure 4.5:** Model Data - Bimodal Distribution: **Blue** – Model Signal (Reference), **Red** – Noisy Signal, **Black** – WavPDS Denoised Signal. The model signal was generated from two Gaussian distributions centered at 4 nm and 5 nm and each with a standard deviation of 0.3 nm. White Gaussian noise was added to generate the Noisy signals at SNRs 30, 10 and 3. **A)** Comparison of Noisy and Denoised signals; **B)** Comparison of the reference signal with the denoised signal obtained by WavPDS; **C)** Distance Distributions from Noisy, Denoised, and Reference signal.

### ***4.3.1.3 Short Evolution Time vs. Long Evolution Time***

High SNR at short evolution times can frequently be obtained but at the cost of poor peak resolution and broader widths in the distance distributions. In these cases, not enough time has been allowed for a reasonably complete dipolar oscillation. On the other hand, due to the reduced SNR at longer evolution times, one is also not able to resolve the peaks due to the presence of increased noise. A major challenge in PDS is to obtain a high SNR signal for longer evolution times. We created a model system to illustrate the effectiveness of WavPDS in denoising the lower SNR signal at a longer evolution time. A model was created, which corresponds to a distance distribution containing two close Gaussian peaks between 5.3 nm and 6 nm. As shown in Fig. 4.6, this distribution is recovered from an 8  $\mu$ s dipolar evolution time signal. This signal was then shortened to 3  $\mu$ s, which resulted in a single peak in the  $P(r)$ , showing the resolution loss at short evolution time. Noisy signals were generated at SNR = 30 for 3  $\mu$ s and at SNR = 10 for 8  $\mu$ s, i.e., a factor of 3 increase of noise assuming the same total signal acquisition times. In many experimental cases one often encounters even greater reduction in SNR by doubling the dipolar evolution times due to effects of phase memory decay. As can be seen in Fig. 4.6, the model, noisy, and denoised signals at 3  $\mu$ s yield a single distance distribution. For longer evolution times at SNR = 10, the denoised signal resolves the two peaks, whereas the noisy signal is unable to. The optimal  $\lambda_{TIKR}$ 's for the noisy signals at 3  $\mu$ s and 8  $\mu$ s are 10 and 20, respectively. For the denoised signals, the optimal  $\lambda_{TIKR}$  was obtained directly from the L-curve and was 1 and 0.01, respectively, for 3  $\mu$ s and 8  $\mu$ s evolution time. The optimal  $\lambda_{TIKR}$ 's for the model signal are  $4.2 \times 10^{-6}$  and  $2.7 \times 10^{-6}$  for the 3  $\mu$ s and 8  $\mu$ s

evolution times, respectively. As mentioned earlier, the near-zero  $\lambda_{TIKR}$  values are due to computational error and are used to eliminate the contributions of the zero singular values due to the singular nature of the matrix  $K$ .



**Figure 4.6:** Model Data - Results of longer evolution time at higher noise (low SNR) compared to shorter evolution time at lower noise (high SNR). **Blue** – Model Signal (Reference), **Red** – Noisy Signal, **Black** – WavPDS Denoised Signal. **A)** Model signal at evolution time  $8 \mu\text{s}$  generated from bimodal distance distribution; **B)** Bimodal distance distribution generated from two Gaussian distributions centered at 5.3 nm and 6 nm, each with standard deviation 0.3 nm, and peak heights of 0.8 and 1, respectively; **C)** Comparison of Noisy, Denoised, and Model signals at  $3 \mu\text{s}$  evolution time for the bimodal distance distribution; **D)** Distance Distributions from Noisy, Denoised, and Model signals at  $3 \mu\text{s}$  evolution time; **E)** Comparison of Noisy, Denoised, and Model signals at  $8 \mu\text{s}$  evolution time for the bimodal distance distribution; **F)** Distance Distributions from Noisy, Denoised, and Model signals at  $8 \mu\text{s}$  evolution time.

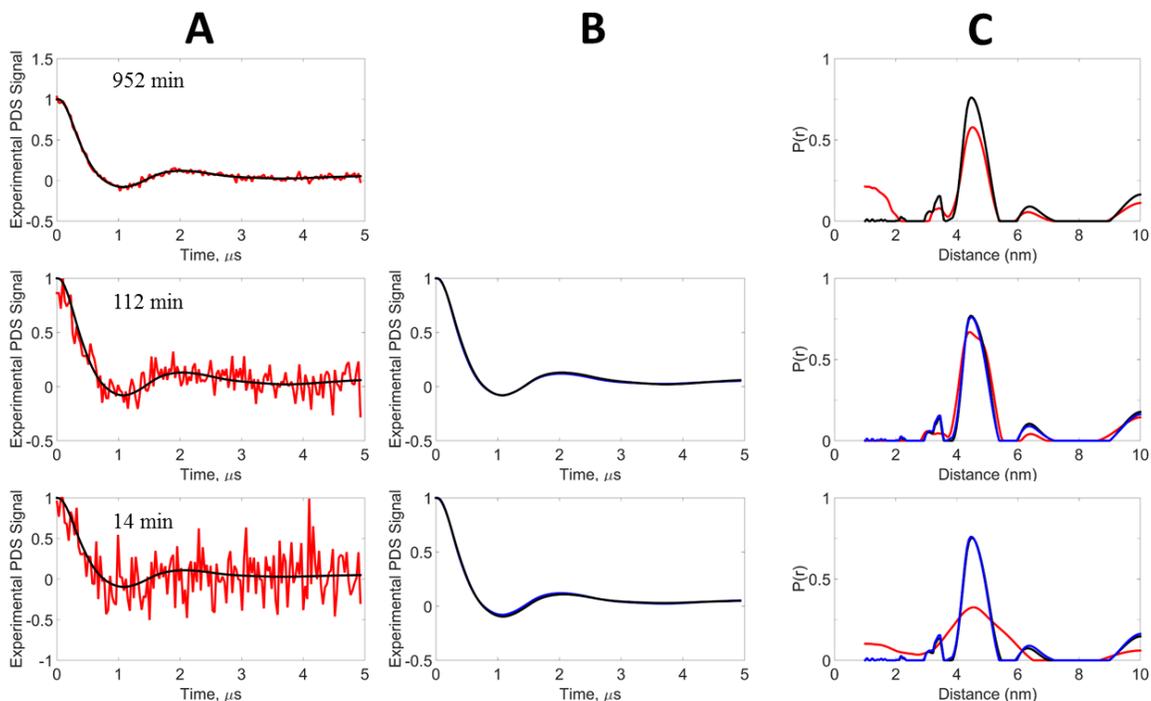
## 4.3.2 Experimental Data

### 4.3.2.1 Unimodal Case

To test WavPDS in experimental signals, the sample was signal-averaged for three different acquisition periods generating small, medium, and large SNRs. Denoising was applied after the baseline was subtracted from the original signal. This was done for simplicity of presentation in this paper. However, we illustrate in Fig. 4.S10 how performing the denoising before baseline subtraction greatly aids in its removal for noisy signals, a matter we plan to discuss elsewhere. This unimodal signal is illustrated in Fig. 4.7A where the DEER signal is averaged for 14 min, 112 min, and 952 min on the protein sample. At 952 min excellent SNR is achieved in the DEER signal (cf. Table 4.3), although some residual noise still affects the longer-time, lower-SNR part of the DEER signal. The  $P(r)$ 's shown in Fig. 4.7C were obtained as in the model cases. A mild sharpening up of the  $P(r)$  is observed after applying WavPDS, wherein most of the residual noise has been removed in the denoised time-domain signal. It was not practical to average for acquisition times much greater than ca. 1000 min. Given the good signal obtained after 952 min we chose to use its denoised form as the reference to compare with the cases of reduced acquisition time. When the signal acquisition time is reduced by almost an order of magnitude (from 952 to 112 min) the noise increases significantly but our denoised signal and  $P(r)$  are almost identical to the reference (cf. Figs. 4.7A, 4.7B, 4.7C), whereas this is not the case for the result obtained from the noisy data (cf. Table 4.3). The last case involves nearly another order of magnitude reduction in signal averaging time to 14 min. This yields a

very noisy signal that is too noisy for a  $P(r)$  to be recovered using  $\lambda_{TIKR}$ , but the result shown in Fig. 4.7C is for  $\lambda_{TIKR}^{OPT}$  and is clearly too broad. However, after our wavelet denoising procedure we recover very nearly identical results as compared to the reference.

Similar to model data results, the  $\lambda_{TIKR}$  values were optimized by manual adjustment for noisy signals, whereas the optimal  $\lambda_{TIKR}$  for the denoised signals was obtained directly from the L-curves (cf. Fig. 4.S11). The  $\lambda_{TIKR}^{OPT}$  value reduced from the noisy signals to denoised signals:  $\lambda_{TIKR}^{OPT} = 30 \rightarrow 0.08$ ,  $\lambda_{TIKR}^{OPT} = 8 \rightarrow 0.07$ , and  $\lambda_{TIKR}^{OPT} = 3 \rightarrow 0.09$  for signal acquisition times 14 min, 112 min, and 952 min, respectively (Fig. 4.S11). It must be noted that denoising improved the L-curve shape to give a sharp elbow and directly resulted in the optimal  $\lambda_{TIKR}$  values, which was not possible for the noisy signal (cf. Fig. 4.S11).



**Figure 4.7:** Experimental Data - Unimodal Distribution: **Blue** – Reference Signal, **Red** – Noisy Signal, **Black** – WavPDS Denoised Signal. The experimental signal was generated from T4 Lysozyme spin-labeled at mutant 44C/ 135C with 63  $\mu\text{M}$  concentration at acquisition times 952 min, 112 min, and 14 min (cf. Section 2.5). **A)** Comparison of Noisy and Denoised signals; **B)** Comparison of Reference signal with Denoised signals after applying WavPDS; **C)** Distance Distributions from Noisy, Denoised, and Reference signal. Denoised signal at 952 min was used as the Reference.

**Table 4.3:** Comparison of Noisy and WavPDS Denoised signals with respect to Experimental reference signal (952 min WavPDS Denoised signal) - Unimodal Distribution at signal acquisition times 952, 112, and 14 min (SNRs 37, 6.8, and 3.8, respectively; cf. Fig. 4.7).

	<b>SNR</b>	$\chi^2$	<b>SSIM</b>
<b>Noisy → Denoised</b>	37 → (Reference)	0.027 → 0 (Reference)	0.676 → 1 (Reference)
	6.8 → 909	0.190 → 0.0011	0.135 → 0.995
	3.8 → 488	0.263 → 0.0021	0.038 → 0.967

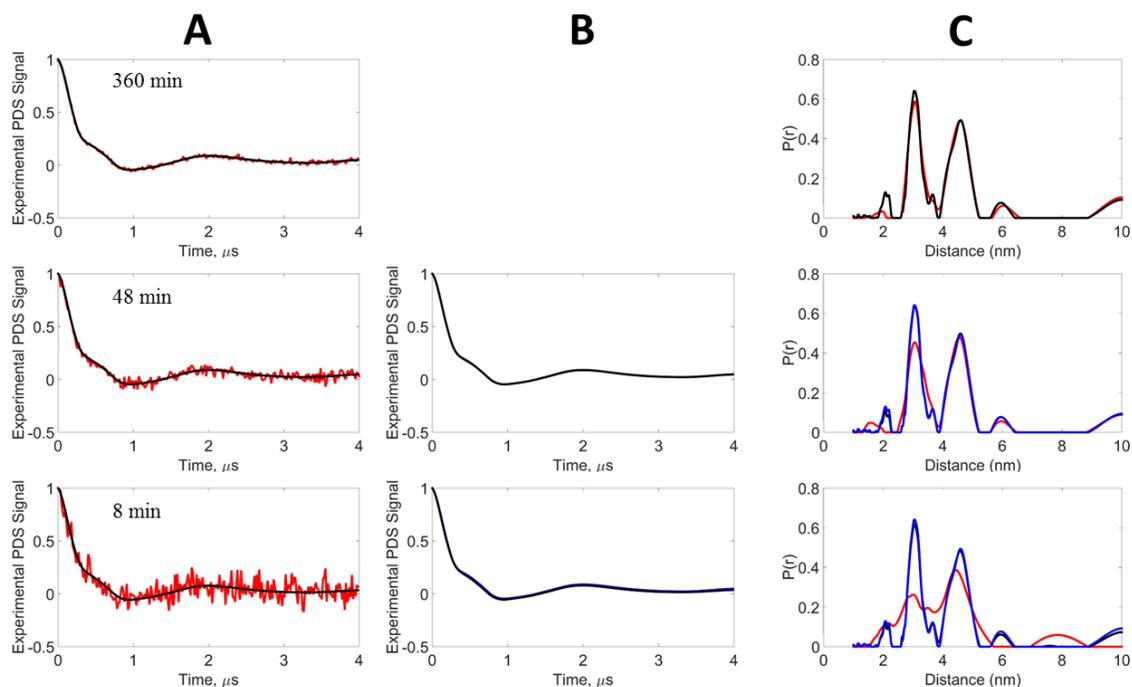
### 4.3.2.2 Bimodal Case

Like the unimodal case, the bimodal signal was obtained at different signal averaging times (8 min, 48 min, and 360 min), generating different SNRs. Fig. 4.8 shows the noisy and denoised signals and their respective distance distributions. It can be seen that at 360 min the signal is nearly noiseless, so the “noisy” and denoised  $P(r)$  are very similar. As in the unimodal case, the denoised 360 min case is taken as the reference for the bimodal signals. For the 48 min noisy signal, the distance distribution is able to retrieve the two peaks but the first peak is now smaller and broader than the second peak, whereas the denoised signal is able to perfectly retrieve both peaks. For 8 min, the noisy signal yields a very broad  $P(r)$  that is barely able to reveal the existence of two peaks. On the other hand, the denoised signal at 8 min is successfully able to retrieve the two peaks, and it is very nearly identical to the reference.

Like the unimodal signal, the  $\lambda_{TIKR}$  values had to be manually adjusted for the noisy signals, whereas the optimal  $\lambda_{TIKR}$  values for the denoised signals were obtained directly from the L-curves (Fig. 4.S11). The  $\lambda_{TIKR}^{OPT}$  values decreased from those of the noisy signals to those of the denoised signals as  $\lambda_{TIKR}^{OPT} = 10 \rightarrow 0.09$ ,  $\lambda_{TIKR}^{OPT} = 5 \rightarrow 0.08$ , and  $\lambda_{TIKR}^{OPT} = 2.48 \rightarrow 0.1$  for signal acquisition times 8 min, 48 min, and 360 min, respectively. It must again be noted that denoising improved the L-curve shape to give a sharp elbow and directly resulted in the optimal  $\lambda_{TIKR}$  values, which was not possible for the noisy signals.

**Table 4.4:** Comparison of Noisy and WavPDS Denoised signals with respect to Experimental reference signal (360 min WavPDS Denoised signal) - Bimodal Distribution at signal acquisition times 360, 48, and 8 min (SNRs 80, 31, and 11, respectively; cf. Fig. 4.8).

	<b>SNR</b>	$\chi^2$	<b>SSIM</b>
<b>Noisy → Denoised</b>	80 → (Reference)	0.013 → 0 (Reference)	0.8815 → 1 (Reference)
	31 → 3333	0.032 → 0.0003	0.573 → 0.999
	11 → 1046	0.086 → 0.0009	0.156 → 0.961



**Figure 4.8:** Experimental Data - Bimodal Distribution: **Blue** – Reference Signal, **Red** – Noisy Signal, **Black** – WavPDS Denoised Signal. The experimental signal was generated from T4 Lysozyme spin-labeled admixture of mutants 8C/44C and 44C/135C at concentrations of 44  $\mu\text{M}$  and 47  $\mu\text{M}$ , respectively, at acquisition times 360 min, 48 min, and 8 min. **A)** Comparison of Noisy and Denoised signals **B)** Comparison of Denoised signals after WavPDS with Reference; **C)** Distance Distributions from Noisy, Denoised, and Reference signal. Denoised signal at 360 min was used as the Reference.

## 4.4 Discussion

This work on denoising time-domain PDS signals, and how this affects the distance distributions that are obtained from them, has led to the following:

- 1) The SNR of the time-domain signal is increased by a factor of about 100 both for model and experimental signals using our new denoising procedure.
- 2) When the  $\text{SNR} \gtrsim 10$ , virtually perfect time-domain signals are obtained after denoising, when compared to a standard; for  $\text{SNR} \sim 3$ , very good denoised signals are obtained.
- 3) Although standard procedures such as SVD followed by TIKR to obtain the  $P(r)$  do a partial denoising by discarding contributions from the heavily noise-corrupted smaller singular values, the prior denoising of the time-domain signals yields significantly better  $P(r)$ 's especially for initially noisy signals where the  $\text{SNR} \lesssim 30$ .
- 4) By denoising the time-domain signals one can employ longer dipolar evolution times, which is important for determining longer distances and/or separating overlapped distance distributions.
- 5) Standard wavelet denoising methods are typically much less effective in denoising noisy time-domain signals compared to WavPDS and hence in obtaining the correct  $P(r)$ , but they become more satisfactory for less noisy signals.

We now discuss these points in more detail.

- 1) Improvement in SNR: We have shown that denoising improves the SNR of

the time-domain signal by a factor of approximately 100 in the model and experimental cases of unimodal and bimodal distributions. In the model cases, we considered noisy SNRs of 3, 10, and 30. The factor of approximately 100 is true in all three SNR scenarios, except that in the unimodal model the SNR increase is about 82. Thus, substantial reductions in signal averaging times can be achieved by denoising. Since signal averaging improves the SNR by  $\sqrt{N}$ , a factor of just 10 in SNR improvement by denoising reduces the signal averaging time by a factor of 100. When we considered SNR  $\sim 1$  for both model cases we found that denoising became unreliable, so we recommend denoising for SNR  $\gtrsim 3$ .

- 2) Accuracy of the denoised signal: Here we find the SSIM parameter best reflects the accuracy of the fit. One needs to compare the time-domain signal to that of a reference to apply this parameter. For the model cases, the original model is the appropriate reference. For the experimental cases it was not feasible to average for long enough periods to obtain ideal references, so we used very long averaging times (ca. 16 and 6 h for the unimodal and bimodal cases, respectively), yielding good SNRs (of 37 and 80, respectively), and then denoised them yielding signals with SNRs of about 3700 and 8000 to serve as the references. For the two model cases for which the SNR = 30, nearly perfect agreement (SSIM of 0.994 and 0.997) with the reference was achieved from the initial noisy SSIM values of 0.58 and 0.54, respectively. When the SNR = 10, almost perfect agreement of 0.96 and 0.95, respectively, was achieved from initial values

for SSIM of  $\sim 0.12, 0.11$ ; while for  $\text{SNR} = 3$  initial values of SSIM 0.01 and 0.008 became 0.95 and 0.945. For the experimental cases for SNRs of 3.8 and 11, SSIM values of 0.97 and 0.96 were achieved after denoising; while for SNRs of 6.8 and 31, SSIM values of 0.995 and 1.00 were achieved. Thus, for an  $\text{SNR} \sim 3-4$ , excellent results can be achieved, while for  $\text{SNR} \sim 7-10$  virtually perfect results are achieved in the time-domain data.

- 3) Improvement in Distance Distributions: Here we consider the implications of denoising the time-domain signals on the  $P(r)$  obtained from them. Although there are other possible approaches to illustrate this, we used TIKR followed by MEM largely to remove negative  $P(r)$  values that emerge from TIKR.<sup>36,51</sup> It is well-known that for  $\text{SNR} \lesssim 30$  this approach, using the L-curve “elbow” to determine the regularization parameters,  $\lambda_{TIKR}^{OPT}$ , is inadequate.<sup>36</sup> We show in Figs. 4.S10 and 4.S11 the L-curve plots. While these plots yielded satisfactory values of  $\lambda$  for the denoised signals, this was not the case for the noisy spectra; we therefore used a trial and error approach to determine the optimum  $\lambda$  for the latter. This was only possible in this present study because we had the reference  $P(r)$ 's for all the cases. In general, determining an optimum  $\lambda$  for noisy cases becomes rather uncertain. But by prior denoising, we find the L-curve works quite well. We show in Figs. 4.4-4.8 how the  $P(r)$ 's derived from the noisy and denoised time-domain signals compare with the respective reference  $P(r)$ 's. Here, because of spurious contributions to the  $P(r)$ 's that appear

(some of which were negative before applying MEM), we did not employ the quantitative measures (i.e.,  $\chi^2$ , SNR, SSIM) that we employed for the time-domain data, but rather relied on visual inspection. It is clear from these figures (Figs. 4.4-4.8) that the denoised time-domain results all yield very good to excellent agreement with the reference  $P(r)$ . This is especially true for initial SNR  $\gtrsim 10$ ; for SNR  $\sim 3$  the results still tend to be very good. The experimental denoised results are particularly good. However, for the noisy spectra we see significantly poorer  $P(r)$  results, especially for SNR  $\sim 3$  and even for SNR  $\sim 10$ , even when using the optimum  $\lambda$ . This is true for both the bimodal model and the experimental results, where the noisy results have difficulty in resolving both peaks. That is, they show broadened  $P(r)$ 's.

Why is depending on just the partial denoising by TIKR less satisfactory than first denoising the time-domain signal before applying TIKR? Since we had to use optimized regularization parameters,  $\lambda_{TIKR}^{OPT}$ , to obtain satisfactory  $P(r)$ 's for the noisy results, that is not the issue. We have indeed found that prior denoising yields significantly smaller  $\lambda$  values than even the optimized values from the noisy signals. These smaller  $\lambda$  values bring into the  $P(r)$  the contributions from smaller singular values in the SVD decomposition of  $P(r)$  (cf. Eq. 4.7) because they are no longer corrupted by noise, and they provide greater resolution (e.g., narrower and more accurate  $P(r)$ 's) and detail that is important for  $P(r)$ 's with more complex structure (e.g., separation of two close distances).

- 4) Longer Distances May be Studied: In the model example shown in Fig. 4.6 we illustrate that when one needs longer dipolar evolution times to distinguish subtle features in the  $P(r)$ , but one is limited by the increase in noise, the denoising method can help to overcome this problem. This is clearly also true when longer dipolar evolution times are needed for the measurement of longer distances.
- 5) Comparison of WavPDS with Standard Methods: Results from the use of Standard Wavelet Denoising Methods<sup>58,59</sup> are shown in the supplement Figs. 4.S3-4.S8 and Tables 4.S1-4.S6 using the same “db6” wavelets. They cover six of the examples: the model bimodal case (Figs. 4.S3 and 4.S4; Tables 4.S1 and 4.S2); and the experimental unimodal (Figs. 4.S5 and 4.S6; Tables 4.S3 and 4.S4) and bimodal (Figs. 4.S7 and 4.S8; Tables 4.S5 and 4.S6) cases. These include a range of initial SNRs. We find that these other methods are much less effective in improving the SNR, i.e., typically only by a factor of 2-3 (except for cases S1 and S2, where the factor is as great as 5) compared to the factor of order 100 for WavPDS. It appears that they are not as successful in removing the noise coefficients in the Detail components representing high-frequency sub-bands nor do they remove any low-frequency noise present in the final Approximation component. The WavPDS method, as we have seen, is very successful in rendering a good  $P(r)$  in all these cases. For the noisiest cases (model bimodal SNR = 3 and experimental unimodal SNR = 3.8), using the standard methods yield poor fits to the double peak model  $P(r)$  (Fig. 4.S3) and better, but not

good, fits to the unimodal  $P(r)$  (Fig. 4.S5). These cases also demonstrate that a simple unimodal  $P(r)$  is more easily fit with lesser quality data than a more complex  $P(r)$ , such as in the bimodal case. This is because contributions from fewer of the smaller singular values are needed to describe the rather simple unimodal distribution. And according to the requirement  $\sigma_i > \lambda_{TIKR}^{OPT}$  (cf. Eq. 4.7) in order that the contribution from the  $i^{th}$  singular value is included, the larger  $\lambda$ 's reduce the number of these significant contributions. Even for a somewhat higher SNR = 11 for the bimodal experimental result (Fig. 4.S6) and an SNR = 10 (Fig. 4.S4) for the bimodal model, the results from the other methods do not yield satisfactory  $P(r)$ 's. However, where one has less noisy signals (e.g., Fig. 4.S8), these other methods do yield more satisfactory  $P(r)$ 's, although not as good as WavPDS.

Low-pass filters are frequently used in denoising, but it is well-known that the standard wavelet denoising methods are superior.<sup>26,29,56</sup> We demonstrate this fact with an example from this work in Fig. 4.S12.

Overall we can conclude that WavPDS is much more effective at denoising the time-domain signals, but as the SNR of the original noisy signal increases and/or the “simplicity” of the actual  $P(r)$  increases, the results from these other methods become more satisfactory. However, this last comment is only true because instead of using the  $\lambda$ -value obtained directly from the elbow of the L-curves, which we find are still poorly defined, we used trial-and-error searches to obtain the optimum  $\lambda$  (which is always

considerably different; cf. Tables 4.S1-4.S6). In this study we have a reference  $P(r)$  for optimizing the choice of  $\lambda$ , but in actual practice, where the true  $P(r)$  is not known a priori, the determination of a good  $\lambda$  still remains a challenge. As we have shown, however, after denoising by WavPDS the L-curve-obtained  $\lambda$  is indeed at or near the optimum  $\lambda$ .

We now wish to note some additional aspects related to the methodology.

- a) Spin-Echo Detection: The experimental time-domain decay signals discussed above are actually obtained as a series of points which themselves represent spin-echo heights obtained as a function of the dipolar evolution time but with the (refocused) spin-echo kept at a fixed time after the initial  $\pi/2$  pulse.<sup>4,67</sup> Each such point is the average of many spin-echo repeats. One can in fact first denoise each of these averaged spin-echo signals. We show an example in Fig. 4.S13 how the spin-echo signal can be very successfully denoised. We plan to investigate this approach further.
- b) Baseline Subtraction from Weak PDS Signals: This is a standard issue in processing PDS data. In the present study we have chosen to first remove the baseline from the original noisy data using our standard procedure<sup>3,68</sup> before denoising, largely to simplify the presentation of the denoising methodology and the comparisons between the various sets of results. Indeed, prior denoising of the raw data, which does include the baseline, would aid significantly in identifying and removing it. This is another matter we plan to address elsewhere, but we do show a preliminary example of this in Fig. 4.S9.

This is an example of a very weak PDS signal from sub- $\mu\text{M}$  concentration of labeled protein. Nevertheless, the denoising is very effective in enabling the removal of the baseline and recovering the denoised signal.

- c) Spurious Peaks in the  $P(r)$ : Even after denoising, spurious peaks still arise in the  $P(r)$ . As we showed in Fig. 4.3, it is not only noise that contributes but truncation of the smaller singular values contributions in the TIKR procedure. This is an effect that remains even after denoising by WavPDS. We plan to study this matter further.
- d) What is the minimum SNR of the original data that can still be usefully denoised by WavPDS? Our results on both model and experimental data show success with an SNR  $\sim 3$ , but further studies are needed to best answer this question.

## 4.5 Conclusion

In this study, we have shown that by greatly enhancing the SNR of time-domain signals, while retaining the fidelity of the underlying signals by means of our new wavelet denoising method, WavPDS, it is possible to reduce signal averaging times by as much as one to two orders of magnitude. This was demonstrated with the use of both model and experimental pulse-dipolar ESR signals. Once denoised, they may be readily inverted by standard methods into distance distributions more conveniently and more accurately than for the original noisy signals.

## 4.S Supplementary Information

### 4.S.1 Introduction

We show in Fig. 4.S1 the Detail and Approximation components of an experimental signal using wavelet “db6.” In Fig. 4.S2, the scaling and wavelet functions of the “db6” wavelet are shown, which are used to generate the Approximation and Detail components, respectively. We show in Figs. 4.S3 to 4.S8 comparisons of some of the WavPDS results from Figs. 4.4, 4.5, 4.7, and 4.8 with the standard methods noted below. These comparisons are also summarized in Tables 4.S1 to 4.S6. In all cases, WavPDS outperforms the other methods. In Fig. 4.S9, we show an example of denoising before baseline subtraction and its utility in detecting baseline.

In Figs. 4.S10 & 4.S11, we show L-curve plots for determining  $\lambda$  both for noisy and denoised data. We distinguish between the  $\lambda_{TIKR}^{L-Curve}$  obtained by locating the point of maximum curvature and those from manual adjustment  $\lambda_{TIKR}^{OPT}$  to best compare with the reference. Whereas they differ for the noisy signals, they are the same for the denoised signals. A comparison of WavPDS and a low-pass filter is shown in Fig. 4.S12. In Fig. 4.S13, the denoised echo signal is shown using this new method.

### 4.S.2 Data Generation Methods

#### 4.S.2.1 Model Signal (Figs. 4.S3-4.S4): Bimodal

The model signal was generated from two Gaussian distributions centered at 4 nm and

5 nm with a standard deviation of 0.3 nm. The peak height of the first peak is 80% of the second peak. White Gaussian noise was added to generate the Noisy signals at SNR = 3 (Fig. 4.S3A, **Red**) and SNR = 10 (Fig. 4.S4A, **Red**).

#### ***4.S.2.2 Experimental Signal (Figs. 4.S5-4.S6): Unimodal***

The experimental signal was generated from T4 Lysozyme spin-labeled at mutant 44C/ 135C with 63  $\mu\text{M}$  concentration (cf. Section 2.5 for details). The signal acquisition time was 14 min (SNR = 3.8, cf. Fig. 4.S5A, **Red**) and 112 min (SNR = 6.8, cf. Fig. 4.S6A, **Red**).

#### ***4.S.2.3 Experimental Signal (Figs. 4.S7-4.S8): Bimodal***

The experimental signal was generated from T4 Lysozyme spin-labeled admixture of mutants 8C/44C and 44C/135C at concentrations of 44  $\mu\text{M}$  and 47  $\mu\text{M}$ , respectively (cf. Section 2.5 for details). The signal acquisition time was 8 min (SNR = 11, cf. Fig. 4.S7A, **Red**) and 48 min (SNR = 31, cf. Fig. 4.S8A, **Red**).

### **4.S.3 Standard Wavelet Denoising Methods**

We compare our WavPDS method with the current standard wavelet denoising methods such as Minimax<sup>59</sup> and SUREShrink<sup>58</sup> for bimodal model signal (Figs. 4.S3-4.S4) and unimodal and bimodal experimental signals (Figs. 4.S5-4.S8) using the same Daubechises “db6” wavelet. They both are used to select optimal noise

thresholds for the Detail components in the wavelet domain. Although there are many new wavelet denoising methods, these two methods are widely used and perform well for all types of signals. We compared our new method with these two methods and others in our previous paper.<sup>25</sup>

#### **4.S.3.1 Noise Thresholding Function**

Both hard and soft thresholding were used for Minimax and SUREShrink method to obtain denoised coefficients. They are referred as Minimax-Hard and SUREShrink-Hard for hard thresholding, and Minimax-Soft and SUREShrink-Soft for soft thresholding in Figs. 4.S3-4.S8. The hard (Eq. 4.S1) and soft (Eq. 4.S2) thresholding are defined as,

$$D'_j[n] = \begin{cases} 0, & \text{for } |D_j[n]| < \lambda_j \\ D_j[n], & \text{otherwise} \end{cases} \quad (4.S1)$$

and

$$D'_j[n] = \begin{cases} 0, & \text{for } |D_j[n]| < \lambda_j \\ \text{sgn}(D_j[n]) (|D_j[n]| - \lambda_j), & \text{otherwise} \end{cases} \quad (4.S2)$$

where  $D_j[n]$  and  $D'_j[n]$  are the noisy and denoised Detail component, respectively, at the  $j^{\text{th}}$  Decomposition level, and  $\lambda_j$  is the noise threshold selected for the  $j^{\text{th}}$  Detail component using the Minimax or SUREShrink method.

#### ***4.S.3.2 Decomposition Level Selection***

For Minimax and SUREShrink methods, the decomposition level that resulted in the denoised signal with highest SNR was selected. To calculate SNR, model data was used as a reference for model signals and WavPDS denoised data at 952 min (unimodal, Figs. 4.S5-4.S6) and 360 min (bimodal, Figs. 4.S5-4.S6) was used for experimental signals.

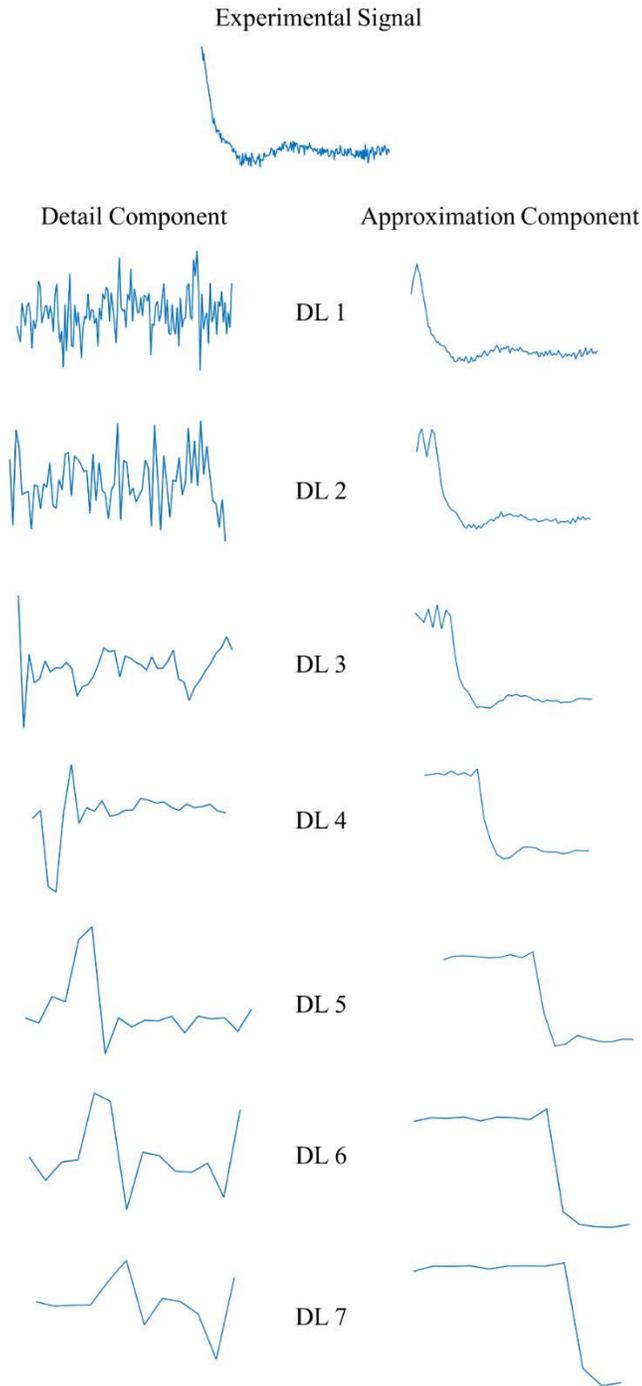
#### ***4.S.3.3 Signal Flipping***

Like WavPDS, the signal  $S(t)$  was flipped to  $S(-t)$  (i.e., time reversed) before applying the standard denoising methods. This is to avoid distorting the initial ( $t = 0$ ) signal as mentioned in Section 2.1. The flipped signal contains the same information as that of the non-flipped signal.

#### **4.S.4 Details of Platform and Software**

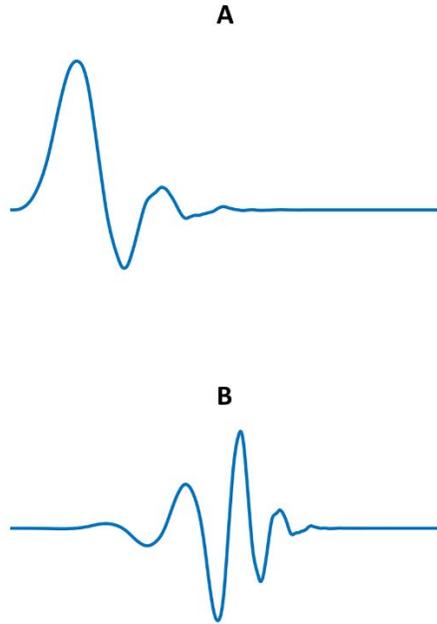
A machine with Intel Core i5-4590 CPU @ 3.30 GHz processor, Windows 7 operating system, 16 GB RAM, and 64-bit operating system was used as the platform. Denoising was performed using MATLAB 2014b. Built-in *wden* was used to denoise signals for the Minimax and SUREShrink wavelet denoising methods. WavPDS codes were written and implemented in MATLAB. Tikhonov Regularization (TIKR)<sup>36</sup> and Maximum Entropy Method (MEM)<sup>51</sup> codes written for MATLAB available on ACERT website ([https://acert.cornell.edu/index\\_files/acert\\_resources.php](https://acert.cornell.edu/index_files/acert_resources.php)) were used to generate  $P(r)$ . For WavPDS and standard methods, Daubechises 6 wavelet (“db6”) was used for Discrete Wavelet Transform (DWT).

### 4.S.5 Wavelet Components of an Experimental PDS Signal



**Figure 4.S1:** The Detail and Approximation components of a PDS signal. DL is the decomposition level.

#### 4.S.6. Scaling and Wavelet functions of the “db6” Wavelet



**Figure 4.S2:** The “db6” wavelet used. A) The Scaling function is used to generate Approximation components (cf. Eq. 4.11); B) The Wavelet function is used to generate Detail components (cf. Eqs. 4.8 and 4.9).

#### 4.S.7 Comparison of Standard Methods with WavPDS

**Table 4.S1:** Comparison of WavPDS and Standard Denoising methods (Minimax and SUREShrink) on Noisy signal for Model Data - Bimodal Distribution at SNR 3 (cf. Fig. 4.S3).

<b>Method</b>	<b>SNR</b>	$\chi^2$	<b>SSIM</b>	$\lambda_{TIKR}^{OPT}$	$\lambda_{TIKR}^{L-Curve}$
Noisy	3	0.333	0.008	60	904
Minimax-Hard	6	0.167	0.282	40	897
Minimax-Soft	19	0.053	0.626	30	76
SUREShrink-Hard	6	0.179	0.119	30	872
SUREShrink-Soft	9	0.107	0.226	30	99
New Method	378	0.0026	0.945	10	10

**Table 4.S2:** Comparison of WavPDS and Standard Denoising methods (Minimax and SUREShrink) on Noisy signal for Model Data - Bimodal Distribution at SNR 10 (cf. Fig. 4.S4).

<b>Method</b>	<b>SNR</b>	$\chi^2$	<b>SSIM</b>	$\lambda_{TIKR}^{OPT}$	$\lambda_{TIKR}^{L-Curve}$
Noisy	10	0.100	0.107	20	85
Minimax-Hard	18	0.054	0.607	15	79
Minimax-Soft	50	0.020	0.871	15	39
SUREShrink-Hard	18	0.054	0.572	15	77
SUREShrink-Soft	38	0.026	0.775	15	51
New Method	1850	0.0005	0.953	3	3

**Table 4.S3:** Comparison of WavPDS and Standard Denoising methods (Minimax and SUREShrink) on Noisy signal for Experimental Data - Unimodal Distribution at signal acquisition time 14 min (SNR 3.8; cf. Fig. 4.S5).

<b>Method</b>	<b>SNR</b>	$\chi^2$	<b>SSIM</b>	$\lambda_{TIKR}^{OPT}$	$\lambda_{TIKR}^{L-Curve}$
Noisy	3.8	0.263	0.038	30	0.08
Minimax-Hard	6	0.160	0.386	10	28
Minimax-Soft	13	0.079	0.306	10	4
SUREShrink-Hard	7	0.142	0.288	10	24
SUREShrink-Soft	11	0.094	0.313	10	26
New Method	488	0.0021	0.967	0.08	0.08

**Table 4.S4:** Comparison of WavPDS and Standard Denoising methods (Minimax and SUREShrink) on Noisy signal for Experimental Data - Unimodal Distribution at signal acquisition time 112 min (SNR 6.8; cf. Fig. 4.S6).

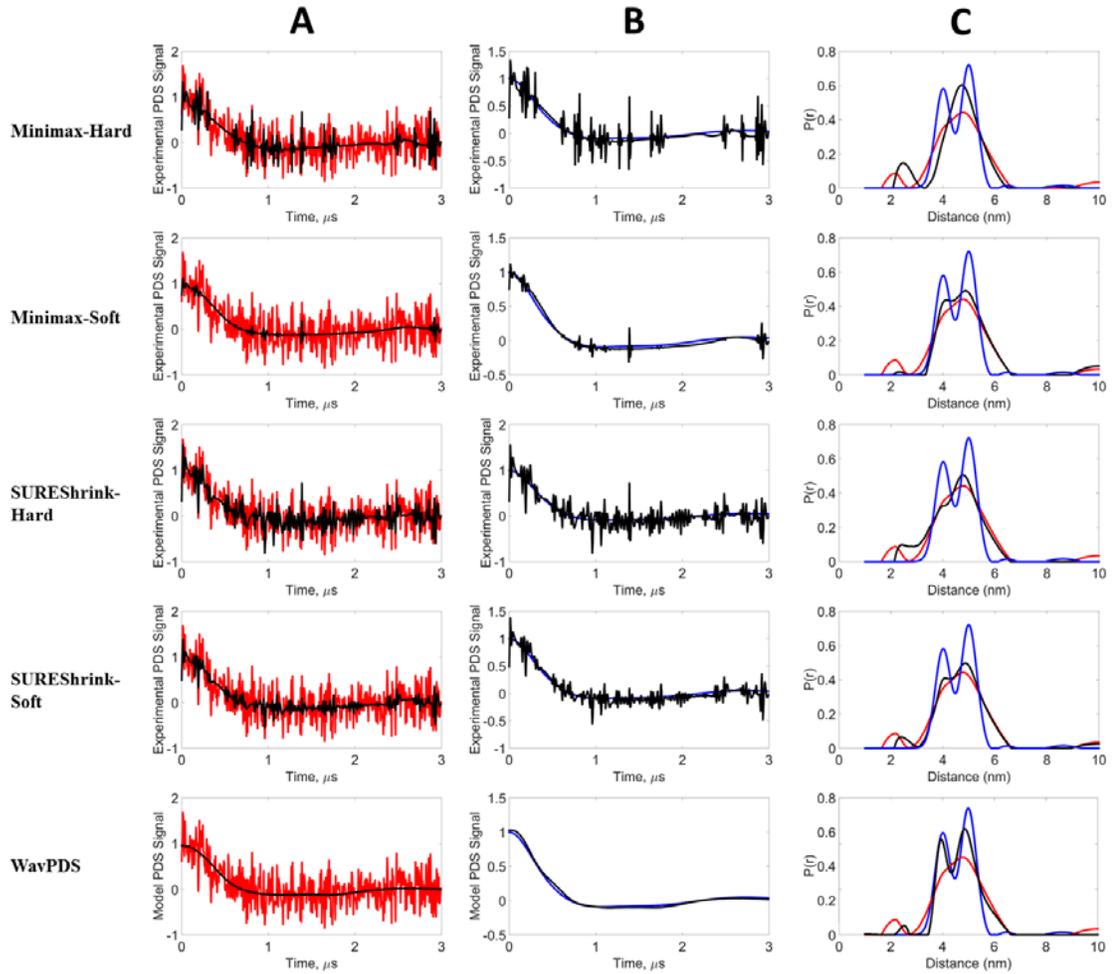
<b>Method</b>	<b>SNR</b>	$\chi^2$	<b>SSIM</b>	$\lambda_{TIKR}^{OPT}$	$\lambda_{TIKR}^{L-Curve}$
Noisy	6.8	0.190	0.135	8	25
Minimax-Hard	12	0.085	0.364	5	0.07
Minimax-Soft	23	0.043	0.486	5	3
SUREShrink-Hard	10	0.098	0.177	5	0.08
SUREShrink-Soft	17	0.058	0.442	5	4
New Method	909	0.0011	0.995	0.07	0.07

**Table 4.S5:** Comparison of WavPDS and Standard Denoising methods (Minimax and SUREShrink) on Noisy signal for Experimental Data - Bimodal Distribution at signal acquisition time 8 min (SNR 11; cf. Fig. 4.S7).

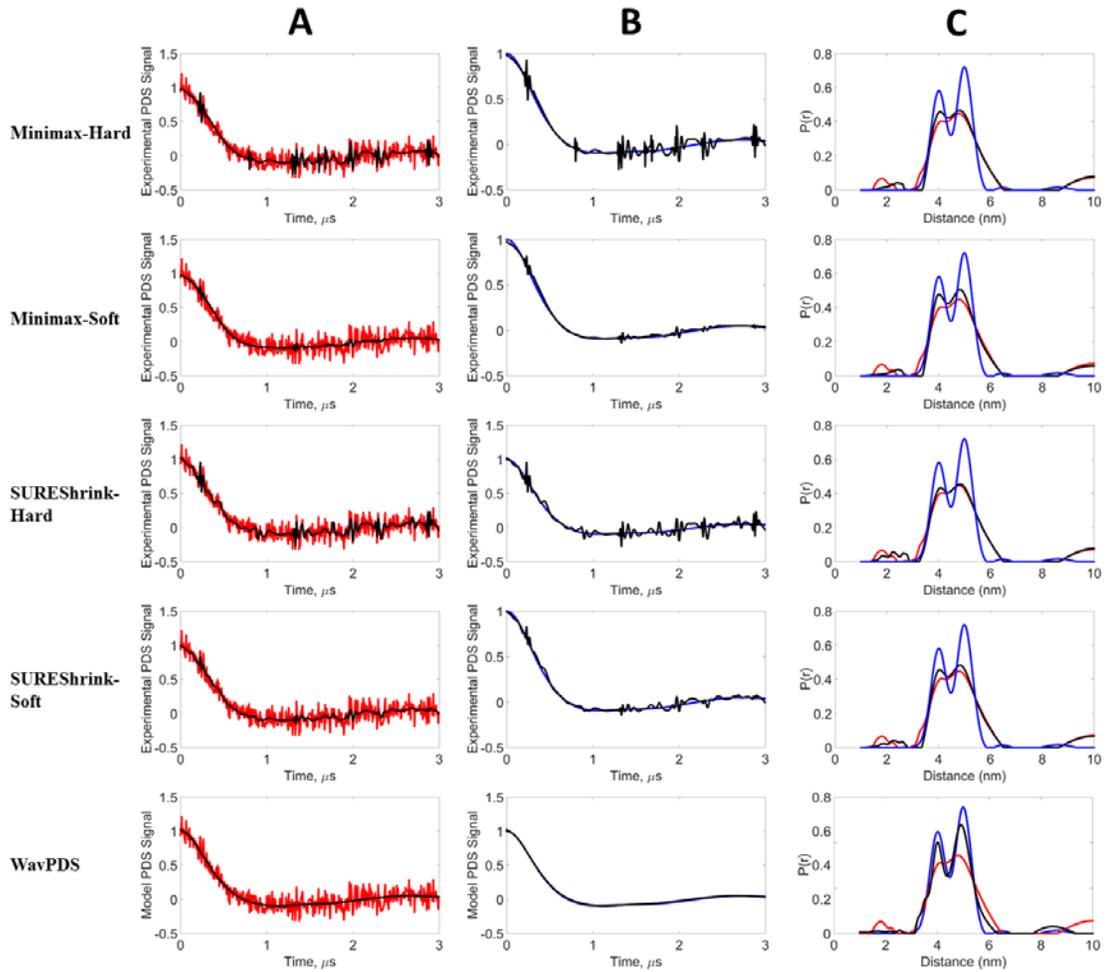
<b>Method</b>	<b>SNR</b>	$\chi^2$	<b>SSIM</b>	$\lambda_{TIKR}^{OPT}$	$\lambda_{TIKR}^{L-Curve}$
Noisy	11	0.086	0.156	10	255
Minimax-Hard	19	0.052	0.418	8	46
Minimax-Soft	33	0.030	0.708	8	31
SUREShrink-Hard	15	0.065	0.266	8	119
SUREShrink-Soft	23	0.044	0.469	8	40
New Method	1046	0.0009	0.961	0.09	0.09

**Table 4.S6:** Comparison of WavPDS and Standard Denoising methods (Minimax and SUREShrink) on Noisy signal for Experimental Data - Bimodal Distribution at signal acquisition time 48 min (SNR 31; cf. Fig. 4.S8).

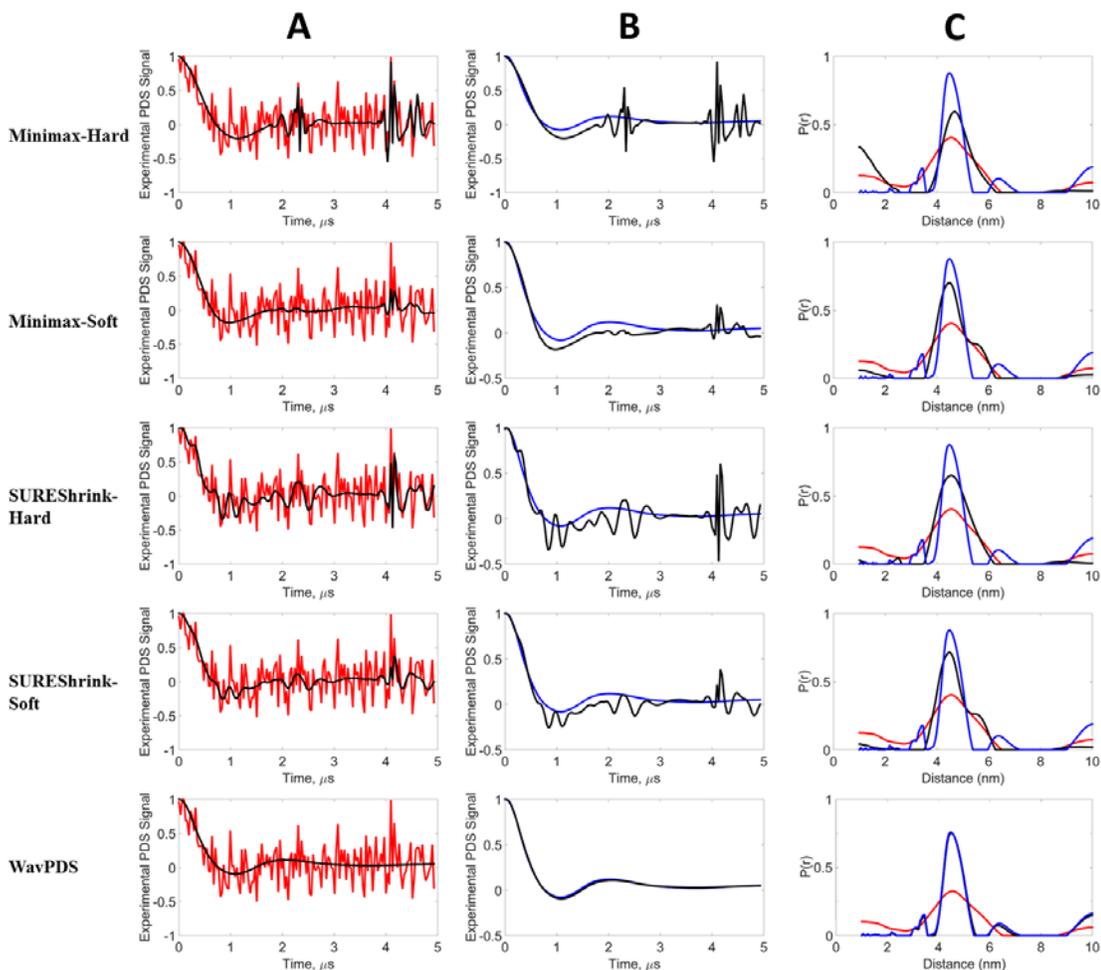
<b>Method</b>	<b>SNR</b>	$\chi^2$	<b>SSIM</b>	$\lambda_{TIKR}^{OPT}$	$\lambda_{TIKR}^{L-Curve}$
Noisy	31	0.032	0.573	5	118
Minimax-Hard	49	0.021	0.798	1	3
Minimax-Soft	89	0.011	0.956	5	0.78
SUREShrink-Hard	44	0.023	0.772	3	45
SUREShrink-Soft	84	0.012	0.930	2.26	2
New Method	3333	0.0003	0.999	0.08	0.08



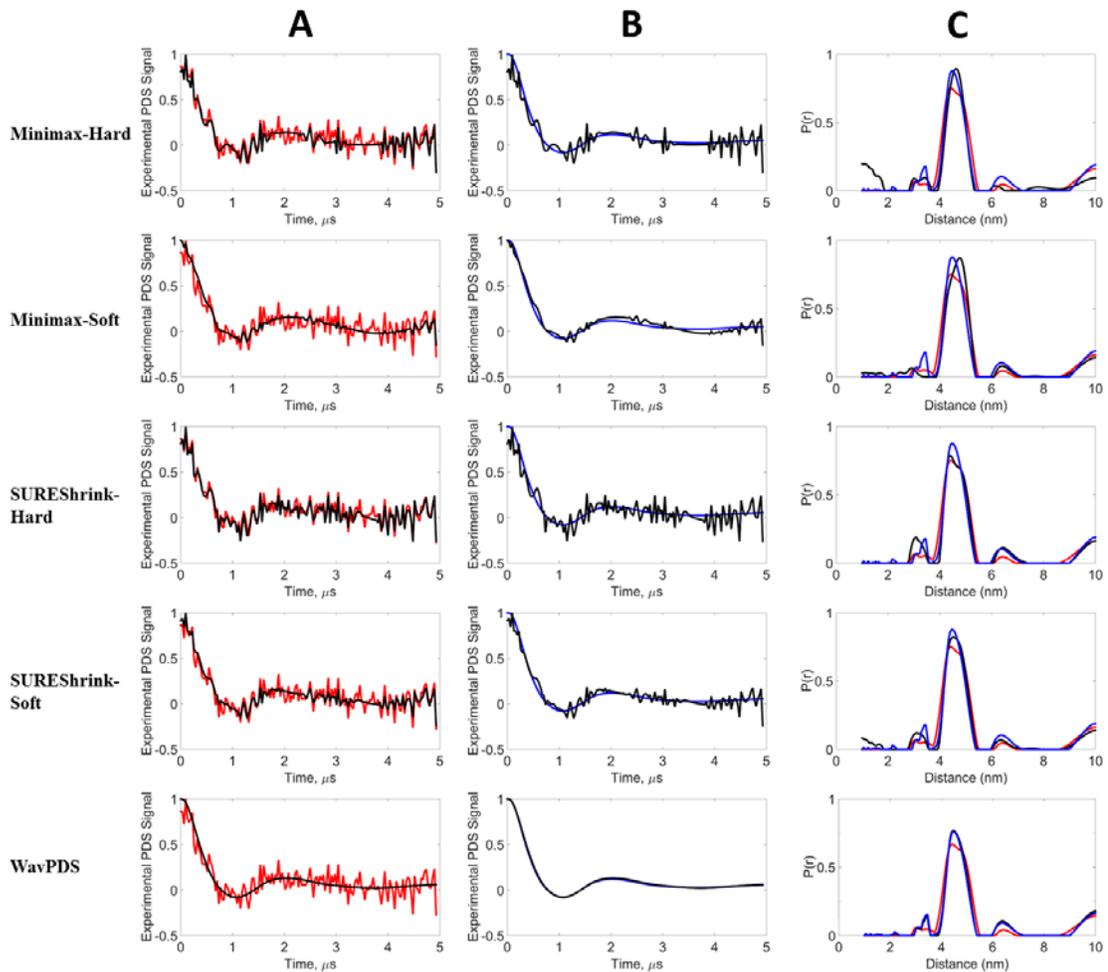
**Figure S3:** Model Data - Bimodal Distribution (cf. Fig. 4.5) with  $\text{SNR} = 3$ : Results of new WavPDS method compared to standard wavelet denoising methods such as Minimax and SUREShrink using hard and soft noise thresholding. **Blue** – Model Signal (Reference), **Red** – Noisy Signal, **Black** – Denoised Signal. **A)** Comparison of Noisy and Denoised signals; **B)** Comparison of Model signals and Denoised signals; **C)** Distance Distributions from Noisy, Denoised, and Model signals.



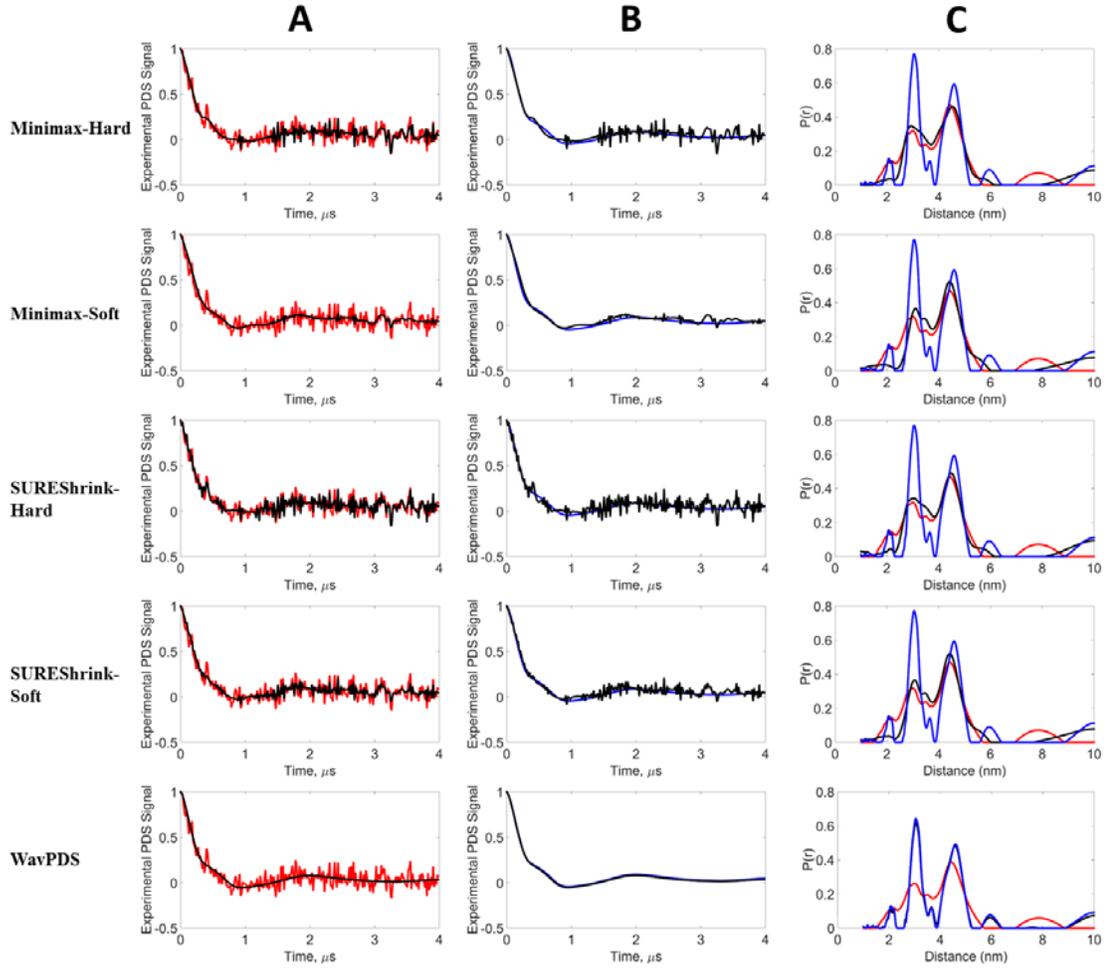
**Figure S4:** Model Data - Bimodal Distribution (cf. Fig. 4.5) with  $\text{SNR} = 10$ : Results of new WavPDS method compared to standard wavelet denoising methods such as Minimax and SUREShrink using hard and soft noise thresholding. **Blue** – Model Signal (Reference), **Red** – Noisy Signal, **Black** – Denoised Signal. **A)** Comparison of Noisy and Denoised signals; **B)** Comparison of Model signals and Denoised signals; **C)** Distance Distributions from Noisy, Denoised, and Model signals.



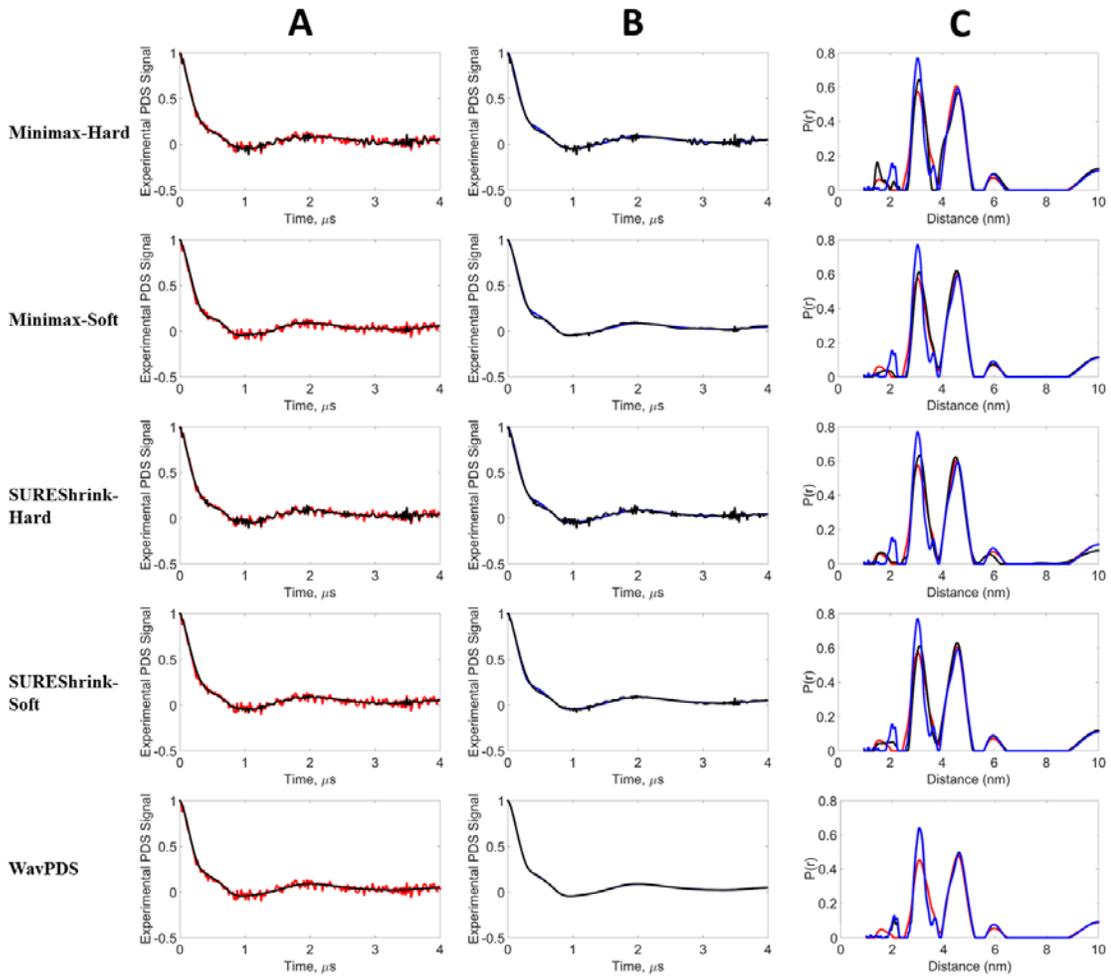
**Figure S5:** Experimental Data - Unimodal Distribution (cf. Fig. 4.7) with 14 min of acquisition time (SNR = 3.8): Results of WavPDS method compared to standard wavelet denoising methods such as Minimax and SUREShrink using hard and soft noise thresholding. **Blue** – Model Signal (Reference), **Red** – Noisy Signal, **Black** – Denoised Signal. **A)** Comparison of Noisy and Denoised signals; **B)** Comparison of Reference signals and Denoised signals; **C)** Distance Distributions from Noisy, Denoised, and Reference signals. Denoised signal at 952 min was used as the Reference.



**Figure S6:** Experimental Data - Unimodal Distribution (cf. Fig. 4.7) with 112 min of acquisition time (SNR = 6.8): Results of new WavPDS method compared to standard wavelet denoising methods such as Minimax and SUREShrink using hard and soft noise thresholding. **Blue** – Reference Signal, **Red** – Noisy Signal, **Black** – Denoised Signal. **A)** Comparison of Noisy and Denoised signals; **B)** Comparison of Reference signals and Denoised signals; **C)** Distance Distributions from Noisy, Denoised, and Reference signals. Denoised signal at 952 min was used as the Reference.

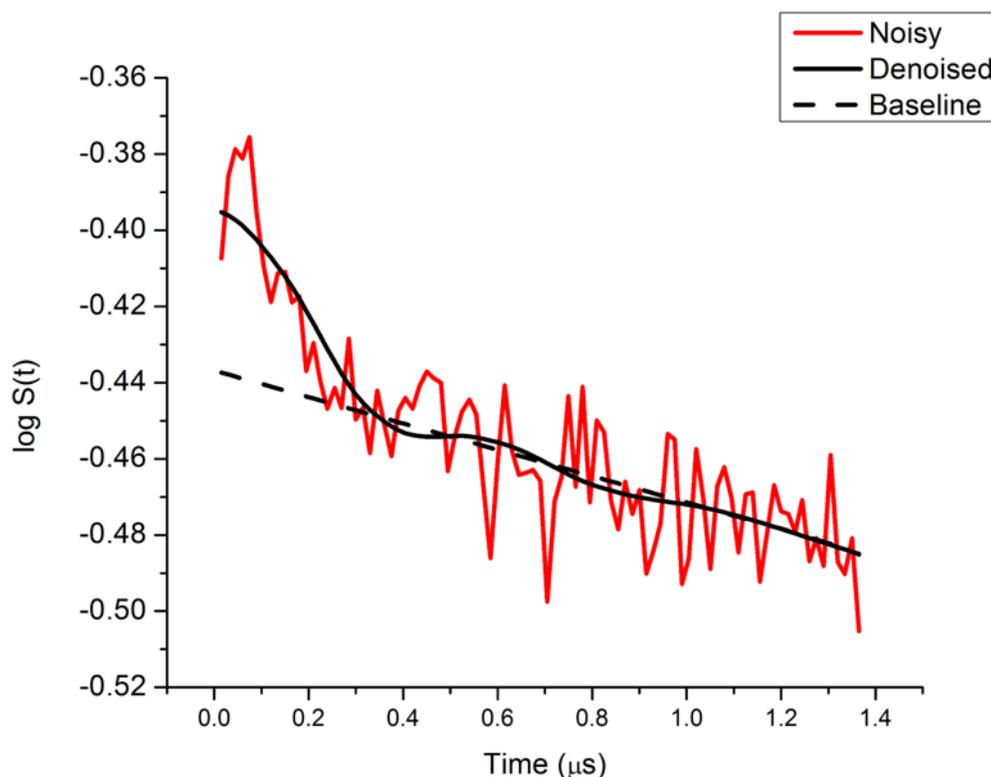


**Figure S7:** Experimental Data - Bimodal Distribution (cf. Fig. 4.8) with 8 min of acquisition time (SNR = 11): Results of new WavPDS method compared to standard wavelet denoising methods such as Minimax and SUREShrink using hard and soft noise thresholding. **Blue** – Reference Signal, **Red** – Noisy Signal, **Black** – Denoised Signal. **A)** Comparison of Noisy and Denoised signals; **B)** Comparison of Reference signals and Denoised signals; **C)** Distance Distributions from Noisy, Denoised, and Reference signals. Denoised signal at 360 min was used as the Reference.



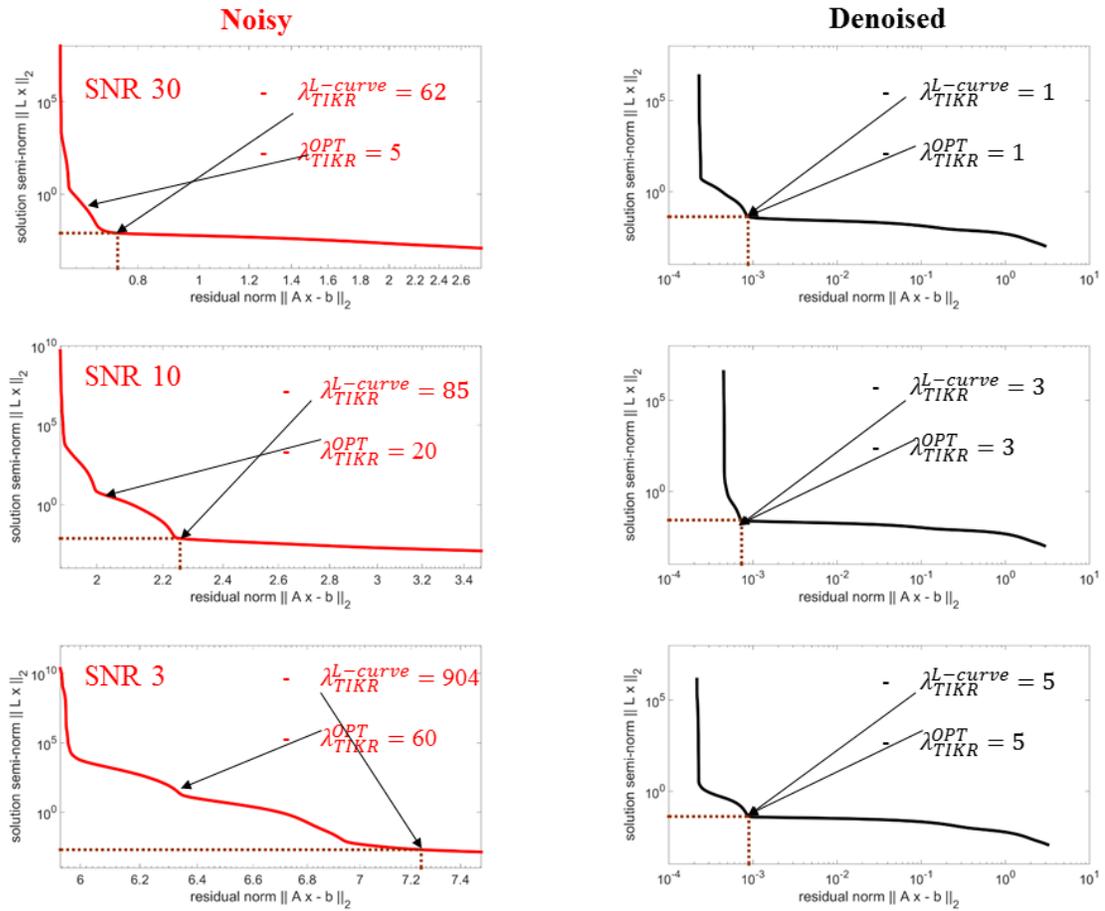
**Figure S8:** Experimental Data - Bimodal Distribution (cf. Fig. 4.8) with 48 min of acquisition time (SNR = 31): Results of new WavPDS method compared to standard wavelet denoising methods such as Minimax and SUREShrink using hard and soft noise thresholding. **Blue** – Reference Signal, **Red** – Noisy Signal, **Black** – Denoised Signal. **A)** Comparison of Noisy and Denoised signals; **B)** Comparison of Reference signals and Denoised signals; **C)** Distance Distributions from Noisy, Denoised, and Reference signals. Denoised signal at 360 min was used as the Reference.

#### 4.S.8 Example of Signal Denoising Before Baseline Subtraction

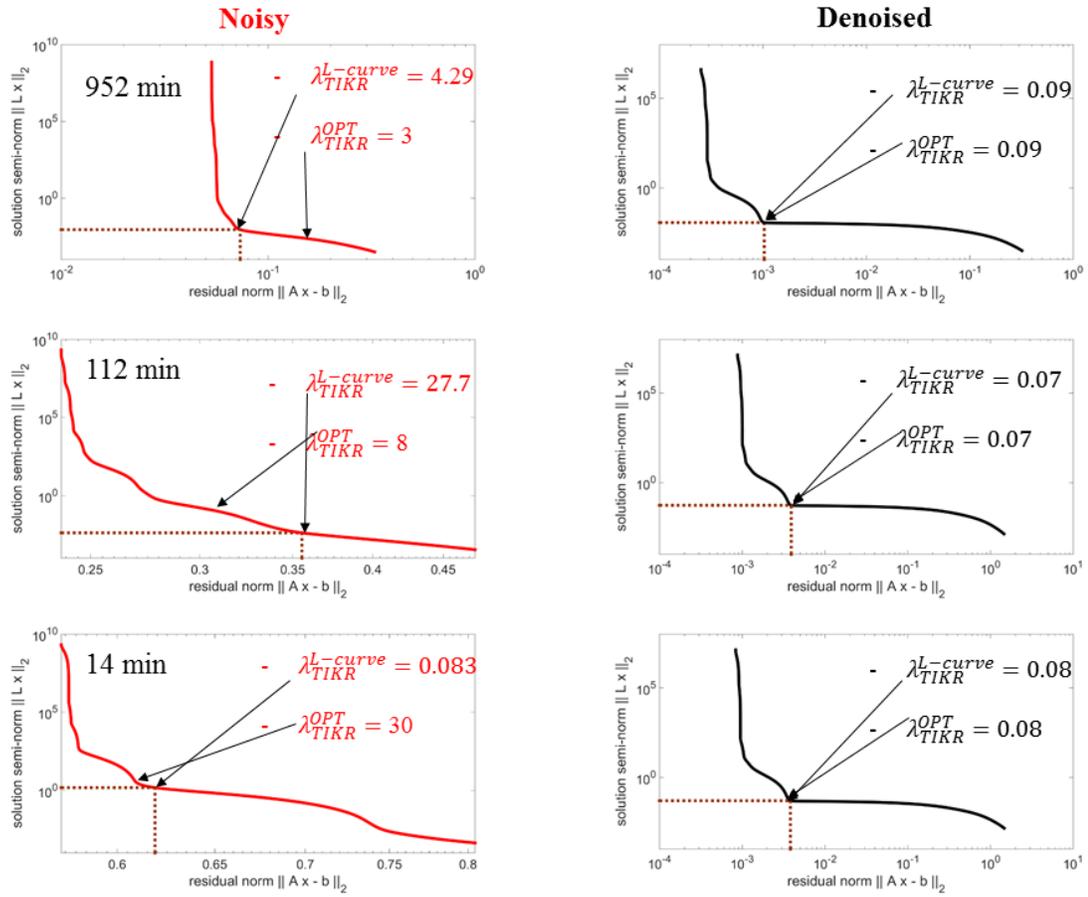


**Figure S9:** Comparison of Noisy signal (after 18 hr of signal averaging) and WavPDS Denoised signal before baseline subtraction. Baseline is determined by fitting the last several points of the denoised spectrum by a straight line. The  $\log$  of the pulsed dipolar signal ( $\log(S(t))$ ) is plotted versus the evolution time ( $\mu\text{s}$ ) for a sub- $\mu\text{M}$  concentration of spin labeled IgE cross-linked with trivalent DNA-DNP ligand in PBS buffer solution.<sup>69</sup> (After baseline removal the SNR is 3.8.) This figure was provided by Siddarth Chandrasekaran (ACERT).

#### 4.S.9 Use of L-curve to Determine $\lambda$

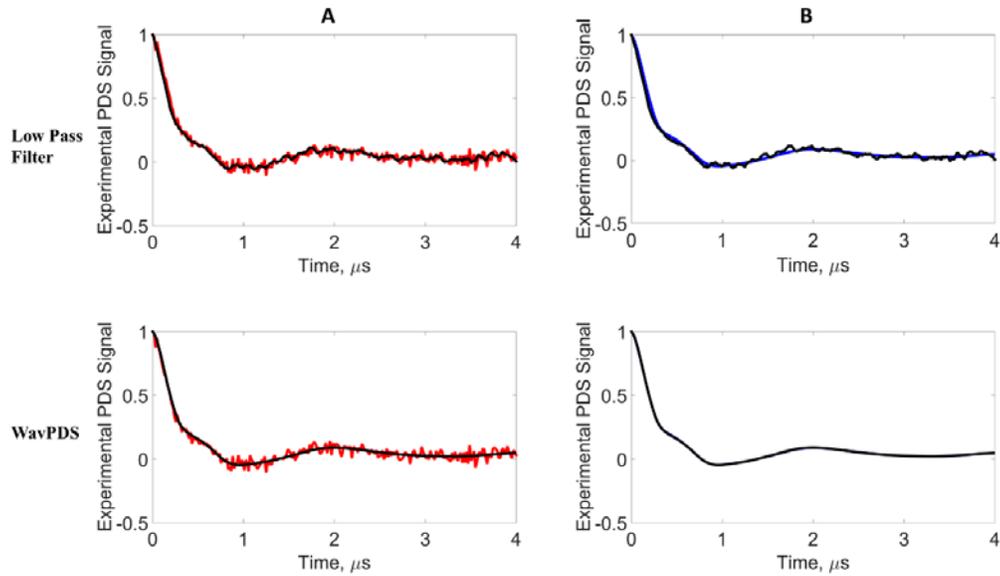


**Figure S10:** L-curve plots to determine  $\lambda$  for Noisy signals at SNRs 30, 10, and 3, and L-curve plots of their respective WavPDS Denoised signals for Model Data - Bimodal Distribution (as shown in Fig. 4.5). **Red** – Noisy Signal L-curve plots, **Black** – Denoised L-curve plots.



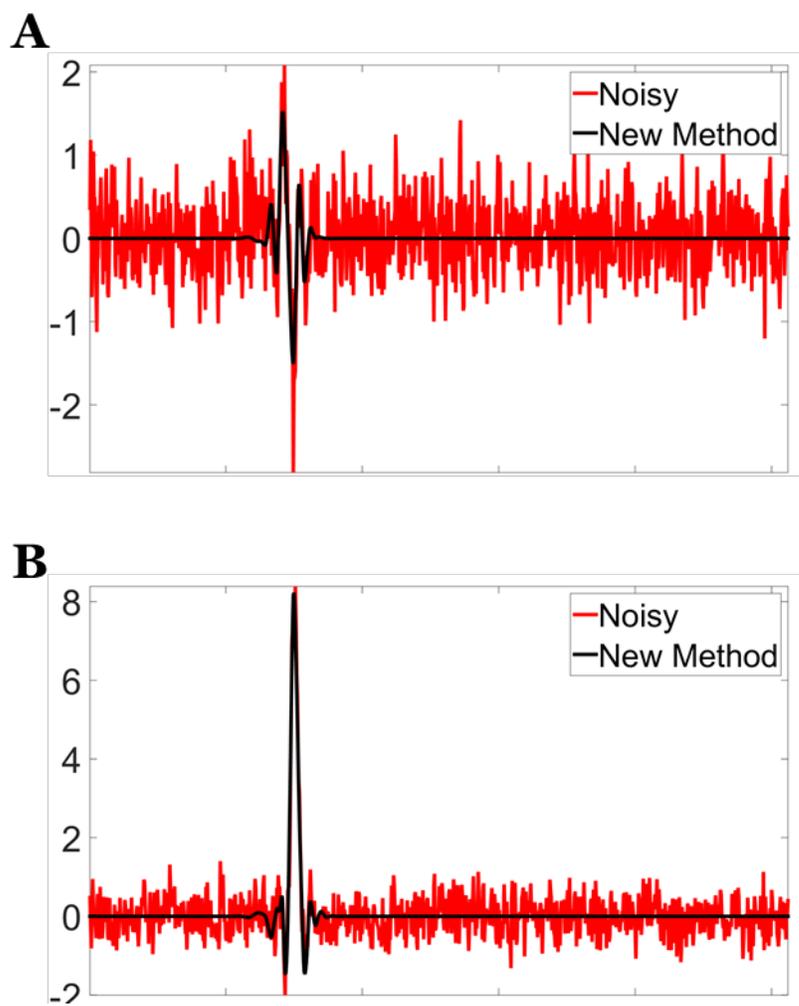
**Figure S11:** L-curve plot to determine  $\lambda$  for Noisy signal at signal acquisition times 952, 112, and 14 min, and the L-curve plot of their respective WavPDS Denoised signals for Experimental Data - Unimodal Distribution (cf. Fig. 4.7). **Red** –Noisy Signal L-curve plots, **Black** – Denoised L-curve plots.

#### 4.S.10 Comparison of a Low-pass Filter with WavPDS



**Figure S12:** Experimental Data—Bimodal Distribution (cf. Fig. 4.8) with 48 min of acquisition time (SNR = 31): Results of WavPDS method compared to low-pass filtering. **Blue**—Reference signal; **Red**—Noisy signal; **Black**—Denoised signal. A) Comparison of the Noisy and Denoised Signals; B) Comparison of Reference and Denoised signals. Denoised signal at 360 min was used as the Reference.

#### 4.S.11 Example of Spin Echo Denoising



**Figure S13:** Echo Signal Denoising: **Red** - Original Noisy Signal, **Black** - Denoised Signal: **A)** Dispersion where SNR goes from 4 to  $1.7 \times 10^7$  by denoising; **B)** Absorption where SNR of 19 becomes  $9.0 \times 10^6$ . The sample was a  $20 \mu\text{M}$  solution of biradical *piperidinyl* -  $\text{CO}_2$  -  $(\text{phenyl})_4$  -  $\text{O}_2\text{C}$  - *piperidinyl*<sup>67</sup> and the  $\pi/2$  and  $\pi/2$  pulses were 2 ns providing full coverage. Coiflet 3 wavelet was used. This figure was provided by Dr. Peter Borbat (ACERT).

## References

- (1) Borbat, P. P.; Freed, J. H. Measuring Distances by Pulsed Dipolar ESR Spectroscopy: Spin-Labeled Histidine Kinases. *Methods Enzymol.* **2007**, *423*, 52–116.
- (2) Borbat, P. P.; Freed, J. H. Pulse Dipolar Electron Spin Resonance: Distance Measurements. In *Structure and Bonding*; Springer Berlin Heidelberg, 2013; pp 1–82.
- (3) Georgieva, E. R.; Borbat, P. P.; Ginter, C.; Freed, J. H.; Boudker, O. Conformational Ensemble of the Sodium-Coupled Aspartate Transporter. *Nat. Struct. Mol. Biol.* **2013**, *20* (2), 215–221.
- (4) Jeschke, G. DEER Distance Measurements on Proteins. *Annu. Rev. Phys. Chem.* **2012**, *63* (1), 419–446.
- (5) Klare, J. P.; Steinhoff, H.-J. Structural Information from Spin-Labelled Membrane-Bound Proteins; Springer Berlin Heidelberg, 2013; pp 205–248.
- (6) Mchaourab, H. S.; Steed, P. R.; Kazmier, K. Toward the Fourth Dimension of Membrane Protein Structure: Insight into Dynamics from Spin-Labeling EPR Spectroscopy. *Structure* **2011**, *19* (11), 1549–1561.
- (7) Edwards, D. T.; Huber, T.; Hussain, S.; Stone, K. M.; Kinnebrew, M.; Kaminker, I.; Matalon, E.; Sherwin, M. S.; Goldfarb, D.; Han, S. Determining the Oligomeric Structure of Proteorhodopsin by Gd<sup>3+</sup>-Based Pulsed Dipolar Spectroscopy of Multiple Distances. *Structure* **2014**, *22* (11), 1677–1686.
- (8) DeBerg, H. A.; Bankston, J. R.; Rosenbaum, J. C.; Brzovic, P. S.; Zagotta, W. N.; Stoll, S. Structural Mechanism for the Regulation of HCN Ion Channels by the Accessory Protein TRIP8b. *Structure* **2015**, *23* (4), 734–744.
- (9) Schiemann, O.; Prisner, T. F. Long-Range Distance Determinations in Biomacromolecules by EPR Spectroscopy. *Q. Rev. Biophys.* **2007**, *40* (01), 1.
- (10) Merz, G. E.; Borbat, P. P.; Pratt, A. J.; Getzoff, E. D.; Freed, J. H.; Crane, B. R. Copper-Based Pulsed Dipolar ESR Spectroscopy as a Probe of Protein Conformation Linked to Disease States. *Biophys. J.* **2014**, *107* (7), 1669–1674.

- (11) Bennati, M.; Robblee, J. H.; Mugnaini, V.; Stubbe, J.; Freed, J. H.; Borbat, P. EPR Distance Measurements Support a Model for Long-Range Radical Initiation in E. Coli Ribonucleotide Reductase. *J. Am. Chem. Soc.* **2005**, *127* (43), 15014–15015.
- (12) Koch, C.; Tria, G.; Fielding, A. J.; Brodhun, F.; Valerius, O.; Feussner, K.; Braus, G. H.; Svergun, D. I.; Bennati, M.; Feussner, I. A Structural Model of PpoA Derived from SAXS-Analysis—Implications for Substrate Conversion. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **2013**, *1831* (9), 1449–1457.
- (13) Hubbell, W. L.; Gross, A.; Langen, R.; Lietzow, M. A. Recent Advances in Site-Directed Spin Labeling of Proteins. *Curr. Opin. Struct. Biol.* **1998**, *8* (5), 649–656.
- (14) Bordignon, E. Site-Directed Spin Labeling of Membrane Proteins. In *EPR Spectroscopy*; Springer Berlin Heidelberg, 2011; pp 121–157.
- (15) Georgieva, E. R.; Xiao, S.; Borbat, P. P.; Freed, J. H.; Eliezer, D. Tau Binds to Lipid Membrane Surfaces via Short Amphipathic Helices Located in Its Microtubule-Binding Repeats. *Biophys. J.* **2014**, *107* (6), 1441–1452.
- (16) Chang, Y.-G.; Cohen, S. E.; Phong, C.; Myers, W. K.; Kim, Y.-I.; Tseng, R.; Lin, J.; Zhang, L.; Boyd, J. S.; Lee, Y.; Kang, S.; Lee, D.; Li, S.; Britt, R. D.; Rust, M. J.; Golden, S. S.; LiWang, A. A Protein Fold Switch Joins the Circadian Oscillator to Clock Output in Cyanobacteria. *Science* (80-. ). **2015**, *349* (6245), 324–328.
- (17) Georgieva, E. R.; Ramlall, T. F.; Borbat, P. P.; Freed, J. H.; Eliezer, D. The Lipid-Binding Domain of Wild Type and Mutant Alpha-Synuclein: Compactness and Interconversion between the Broken and Extended Helix Forms. *J. Biol. Chem.* **2010**, *285* (36), 28261–28274.
- (18) Ward, R.; Bowman, A.; Sozudogru, E.; El-Mkami, H.; Owen-Hughes, T.; Norman, D. G. EPR Distance Measurements in Deuterated Proteins. *J. Magn. Reson.* **2010**, *207* (1), 164–167.
- (19) Borbat, P. P.; Crepeau, R. H.; Freed, J. H. Multifrequency Two-Dimensional Fourier Transform ESR: An X/Ku-Band Spectrometer. *J. Magn. Reson.* **1997**, *127* (2), 155–167.

- (20) Polyhach, Y.; Bordignon, E.; Tschaggelar, R.; Gandra, S.; Godt, A.; Jeschke, G. High Sensitivity and Versatility of the DEER Experiment on Nitroxide Radical Pairs at Q-Band Frequencies. *Phys. Chem. Chem. Phys.* **2012**, *14* (30), 10762.
- (21) Reginsson, G. W.; Hunter, R. I.; Cruickshank, P. A. S.; Bolton, D. R.; Sigurdsson, S. T.; Smith, G. M.; Schiemann, O. W-Band PELDOR with 1kW Microwave Power: Molecular Geometry, Flexibility and Exchange Coupling. *J. Magn. Reson.* **2012**, *216*, 175–182.
- (22) Hertel, M. M.; Denysenkov, V. P.; Bennati, M.; Prisner, T. F. Pulsed 180-GHz EPR/ENDOR/PELDOR Spectroscopy. *Magn. Reson. Chem.* **2005**, *43* (S1), S248–S255.
- (23) Borbat, P. P.; Georgieva, E. R.; Freed, J. H. Improved Sensitivity for Long-Distance Measurements in Biomolecules: Five-Pulse Double Electron-Electron Resonance. *J. Phys. Chem. Lett.* **2013**, *4* (1), 170–175.
- (24) Spindler, P. E.; Glaser, S. J.; Skinner, T. E.; Prisner, T. F. Broadband Inversion PELDOR Spectroscopy with Partially Adiabatic Shaped Pulses. *Angew. Chemie Int. Ed.* **2013**, *52* (12), 3425–3429.
- (25) Srivastava, M.; Anderson, C. L.; Freed, J. H. A New Wavelet Denoising Method for Selecting Decomposition Levels and Noise Thresholds. *IEEE Access* **2016**, *4*, 3862–3877.
- (26) Mallat, S. *A Wavelet Tour of Signal Processing: A Sparse Way*, Third Edit.; Academic Press: New York, 2009.
- (27) Donoho, D. L. De-Noising by Soft-Thresholding. *IEEE Trans. Inf. Theory* **1995**, *41* (3), 613–627.
- (28) Daubechies, I. Ten Lectures on Wavelets. *J. Acoust. Soc. Am.* **1993**, *93* (3), 1671.
- (29) Addison, P. S. *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance*; CRC Press: Philadelphia, 2002.
- (30) Daubechies, I. The Wavelet Transform, Time-Frequency Localization and

Signal Analysis. *IEEE Trans. Inf. Theory* **1990**, 36 (5), 961–1005.

- (31) Ernst, R. R.; Bodenhausen, G.; Wokaun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*; Clarendon Press: Oxford, U.K., 1987.
- (32) Hamm, P.; Zanni, M. *Concepts and Methods of 2D Infrared Spectroscopy*, First Edit.; Cambridge University Press: Cambridge, U.K., 2011.
- (33) Parson, W. W. *Modern Optical Spectroscopy*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2015.
- (34) Pake, G. E. Nuclear Resonance Absorption in Hydrated Crystals: Fine Structure of the Proton Line. *J. Chem. Phys.* **1948**, 16 (4), 327.
- (35) Jeschke, G.; Koch, A.; Jonas, U.; Godt, A. Direct Conversion of EPR Dipolar Time Evolution Data to Distance Distributions. *J. Magn. Reson.* **2002**, 155 (1), 72–82.
- (36) Chiang, Y.-W.; Borbat, P. P.; Freed, J. H. The Determination of Pair Distance Distributions by Pulsed ESR Using Tikhonov Regularization. *J. Magn. Reson.* **2005**, 172 (2), 279–295.
- (37) Raitsimring, A.; Salikhov, K. Electron Spin Echo Method as Used to Analyze the Spatial Distribution of Paramagnetic Centers. *Bull Magn Reson* **1985**, 7 (4), 184–217.
- (38) Milov, A. D.; Maryasov, A. G.; Tsvetkov, Y. D. Pulsed Electron Double Resonance (PELDOR) and Its Applications in Free-Radicals Research. *Appl. Magn. Reson.* **1998**, 15 (1), 107–143.
- (39) Fajer, P. G.; Brown, L.; Song, L. Practical Pulsed Dipolar ESR (DEER). In *ESR Spectroscopy in Membrane Biophysics*; Springer US: Boston, MA, 2007; pp 95–128.
- (40) Brandon, S.; Beth, A. H.; Hustedt, E. J. The Global Analysis of DEER Data. *J. Magn. Reson.* **2012**, 218, 93–104.
- (41) Milov, A. D.; Tsvetkov, Y. D.; Formaggio, F.; Oancea, S.; Toniolo, C.; Raap, J. Aggregation of Spin Labeled Trichogin GA IV Dimers: Distance Distribution between Spin Labels in Frozen Solutions by PELDOR Data. *J. Phys. Chem. B*

**2003**, *107* (49), 13719–13727.

- (42) Borbat, P. P.; Mchaourab, H. S.; Freed, J. H. Protein Structure Determination Using Long-Distance Constraints from Double-Quantum Coherence ESR: Study of T4 Lysozyme. *J. Am. Chem. Soc.* **2002**, *124* (19), 5304–5314.
- (43) Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J.; Timmel, C. R.; Hilger, D.; Jung, H. DeerAnalysis2006—a Comprehensive Software Package for Analyzing Pulsed ELDOR Data. *Appl. Magn. Reson.* **2006**, *30* (3–4), 473–498.
- (44) Stein, R. A.; Beth, A. H.; Hustedt, E. J. A Straightforward Approach to the Analysis of Double Electron–Electron Resonance Data. In *Methods in enzymology*; 2015; Vol. 563, pp 531–567.
- (45) Dalmas, O.; Hyde, H. C.; Hulse, R. E.; Perozo, E. Symmetry-Constrained Analysis of Pulsed Double Electron–Electron Resonance (DEER) Spectroscopy Reveals the Dynamic Nature of the KcsA Activation Gate. *J. Am. Chem. Soc.* **2012**, *134* (39), 16360–16369.
- (46) Pannier, M.; Schädler, V.; Schöps, M.; Wiesner, U.; Jeschke, G.; Spiess, H. W. Determination of Ion Cluster Sizes and Cluster-to-Cluster Distances in Ionomers by Four-Pulse Double Electron Electron Resonance Spectroscopy. *Macromolecules* **2000**, *33* (21), 7812–7818.
- (47) Sen, K. I.; Logan, T. M.; Fajer, P. G. Protein Dynamics and Monomer–Monomer Interactions in AntR Activation by Electron Paramagnetic Resonance and Double Electron–Electron Resonance †. *Biochemistry* **2007**, *46* (41), 11639–11649.
- (48) Masureel, M.; Martens, C.; Stein, R. A.; Mishra, S.; Ruyschaert, J.-M.; Mchaourab, H. S.; Govaerts, C. Protonation Drives the Conformational Switch in the Multidrug Transporter LmrP. *Nat. Chem. Biol.* **2013**, *10* (2), 149–155.
- (49) Jeschke, G.; Panek, G.; Godt, A.; Bender, A.; Paulsen, H. Data Analysis Procedures for Pulse ELDOR Measurements of Broad Distance Distributions. *Appl. Magn. Reson.* **2004**, *26* (1–2), 223–244.
- (50) Bowman, M. K.; Maryasov, A. G.; Kim, N.; DeRose, V. J. Visualization of Distance Distribution from Pulsed Double Electron–Electron Resonance Data.

*Appl. Magn. Reson.* **2004**, 26 (1–2), 23–39.

- (51) Chiang, Y.-W.; Borbat, P. P.; Freed, J. H. Maximum Entropy: A Complement to Tikhonov Regularization for Determination of Pair Distance Distributions by Pulsed ESR. *J. Magn. Reson.* **2005**, 177 (2), 184–196.
- (52) Edwards, T. H.; Stoll, S. A Bayesian Approach to Quantifying Uncertainty from Experimental Noise in DEER Spectroscopy. *J. Magn. Reson.* **2016**, 270, 87–97.
- (53) Dzuba, S. A. The Determination of Pair-Distance Distribution by Double Electron–electron Resonance: Regularization by the Length of Distance Discretization with Monte Carlo Calculations. *J. Magn. Reson.* **2016**, 269, 113–119.
- (54) Antoniadis, A.; Oppenheim, G. *Wavelets and Statistics*; Springer Science & Business Media, 2012.
- (55) Donoho, D. L. Statistical Estimation and Optimal Recovery. *Ann. Stat.* **1994**, 22, 238–270.
- (56) Fodor, I. K. Denoising through Wavelet Shrinkage: An Empirical Study. *J. Electron. Imaging* **2003**, 12 (1), 151.
- (57) Daubechies, I. *Ten Lectures on Wavelets*; CBMS-NSF Regional Conference Series in Applied Mathematics; Society for Industrial and Applied Mathematics, 1992.
- (58) Donoho, D. L.; Johnstone, I. M. Adapting to Unknown Smoothness via Wavelet Shrinkage. *J. Am. Stat. Assoc.* **1995**, 90 (432), 1200.
- (59) Donoho, D. L.; Johnstone, J. M. Ideal Spatial Adaptation by Wavelet Shrinkage. *Biometrika* **1994**, 81 (3), 425–455.
- (60) Wang, Z.; Bovik, A.; Sheikh, H. R.; Simoncelli, E. P. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Trans. Image Process.* **2004**, 13 (4), 600–612.
- (61) Zhou Wang; Bovik, A. Mean Squared Error: Love It or Leave It? A New Look at Signal Fidelity Measures. *IEEE Signal Process. Mag.* **2009**, 26 (1), 98–117.

- (62) Georgieva, E. R.; Roy, A. S.; Grigoryants, V. M.; Borbat, P. P.; Earle, K. A.; Scholes, C. P.; Freed, J. H. Effect of Freezing Conditions on Distances and Their Distributions Derived from Double Electron Electron Resonance (DEER): A Study of Doubly-Spin-Labeled T4 Lysozyme. *J. Magn. Reson.* **2012**, *216*, 69–77.
- (63) Pannier, M.; Veit, S.; Godt, A.; Jeschke, G.; Spiess, H. W. Dead-Time Free Measurement of Dipole–dipole Interactions between Electron Spins. *J. Magn. Reson.* **2011**, *213* (2), 316–325.
- (64) Calvetti, D.; Morigi, S.; Reichel, L.; Sgallari, F. Tikhonov Regularization and the L-Curve for Large Discrete Ill-Posed Problems. *J. Comput. Appl. Math.* **2000**, *123* (1), 423–446.
- (65) Hansen, P. C. Analysis of Discrete Ill-Posed Problems by Means of the L-Curve. *SIAM Rev.* **1992**, *34* (4), 561–580.
- (66) Hansen, P. C.; O’Leary, D. P. The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems. *SIAM J. Sci. Comput.* **1993**, *14* (6), 1487–1503.
- (67) Borbat, P. P.; Freed, J. H. Multiple-Quantum ESR and Distance Measurements. *Chem. Phys. Lett.* **1999**, *313* (1–2), 145–154.
- (68) Georgieva, E. R.; Borbat, P. P.; Norman, H. D.; Freed, J. H. Mechanism of Influenza A M2 Transmembrane Domain Assembly in Lipid Membranes. *Sci. Rep.* **2015**, *5*, 11757.
- (69) Sil, D.; Lee, J. B.; Luo, D.; Holowka, D.; Baird, B. Trivalent Ligands with Rigid DNA Spacers Reveal Structural Requirements for IgE Receptor Signaling in RBL Mast Cells. *ACS Chem. Biol.* **2007**, *2* (10), 674–684.

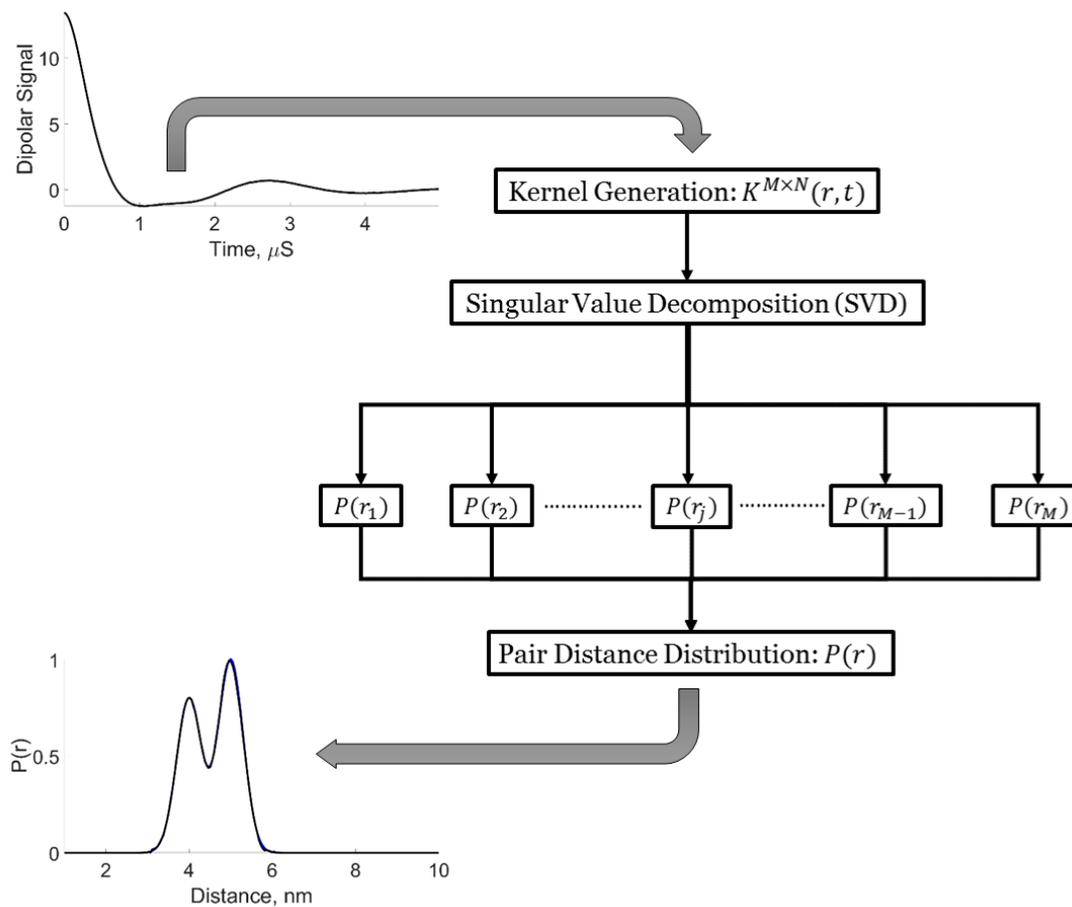
## CHAPTER 5

### SINGULAR VALUE DECOMPOSITION METHOD TO DETERMINE DISTANCE DISTRIBUTIONS IN PULSED DIPOLAR ELECTRON SPIN RESONANCE

#### Abstract

Regularization is often utilized to elicit the desired physical results from experimental data. The recent development of a denoising procedure yielding about two orders of magnitude in improvement in SNR obviates the need for regularization, which achieves a compromise between canceling effects of noise and obtaining an estimate of the desired physical results. We show how singular value decomposition (SVD) can be employed directly on the denoised data, using pulse dipolar electron spin resonance experiments as an example. Such experiments are useful in measuring distances and their distributions,  $P(r)$  between spin labels on proteins. In noise-free model cases exact results are obtained, but even a small amount of noise (e.g.  $SNR = 850$  after denoising) corrupts the solution. We develop criteria that precisely determine an optimum approximate solution, which can readily be automated. This method is applicable to any signal that is currently processed with regularization of its SVD analysis.

## Graphical Abstract



\*This Chapter was published in The Journal of Physical Chemistry Letters

**Authors:** Madhur Srivastava, and Jack H. Freed

**Author Contributions:** Madhur Srivastava conducted the research and wrote the paper, and Jack Freed supervised the research and edited the paper.

## 5.1 Introduction

Many physical techniques require that the original experimental signal be processed to yield the desired results.<sup>1-6</sup> In such cases a regularization method, such as Tikhonov regularization,<sup>7-11</sup> is often employed to strike a balance between canceling undesirable effects of noise in the signal and providing a useful approximation to the desired experimental results. The regularization is usually incorporated with singular-value decomposition (SVD),<sup>7,12</sup> which, however, would be seriously noise-corrupted without regularization.

Here, we describe a new approach which does not depend on the compromise of a regularization method, so it is able to yield significantly more accurate results. We have previously introduced a wavelet-based denoising method which provides SNR improvement of about two orders of magnitude.<sup>13,14</sup> Our new SVD approach reported here benefits from such initial denoising. We illustrate the method by applying it to pulse dipolar ESR (PDS) experiments, which have in recent years been shown to be important in determining distance distributions and structures in many proteins.<sup>2,15-23</sup> But it would be applicable to any signal that currently requires regularization of its SVD analysis.

Our presentation consists of the following aspects: 1) We first show how a careful analysis of the SVD solution yields the exact result for noise-free models. 2) Then we show that even very small amounts of noise (e.g.  $SNR \approx 850$ ) are sufficient to corrupt

this, but we develop a precise criterion to determine which singular value contributions (SVCs) must be kept and which discarded to obtain a useful solution. 3) However, even such a solution typically contains spurious defects, requiring a further refinement of the method that we call *segmentation*. That is, for example, for PDS one obtains a time-dependent signal,  $S(t)$ , from which a distance distribution  $P(r)$  is extracted. It turns out, especially in the presence of finite noise, that the SVD convergence properties are dependent on  $r$ , so that one must apply our method in an  $r$ -dependent fashion. This may be implemented in a practical manner in terms of a few segments of  $r$ . 4) We find that this method leads to a highly accurate  $P(r)$  without spurious peaks, limited only by the residual noise, and it is more accurate than can be achieved using a regularization approach even on the (mostly) denoised signal (cf. Figs. 5.S9-5.S12).

## 5.2 Pulsed Dipolar Spectroscopy

The distance distribution between an ensemble of spin pairs obtained from the PDS signal may be expressed as a Fredholm Equation of the first kind as:

$$\int_{R_{min}}^{R_{max}} \kappa(r, t) P(r) dr = S(t) \tag{5.1}$$

where  $\kappa(r, t)$  is the kernel representing the dipolar interaction of a spin-pair at a given  $r$  that has been integrated over all possible orientations of  $\mathbf{r}$  with respect to the large

magnetic field defining the lab frame;  $P(r)$  is the distance distribution between the spin pairs; and  $S(t)$  is the dipolar signal. Equation 5.1 is discretized in matrix form as follows,

$$KP = S \tag{5.2}$$

where  $K$  is an  $M \times N$  matrix, with  $M$  the dimension of the discrete values of  $S$  and  $N$  of  $P$ , and  $N \leq M$ . The PDS experiment measures the dipolar signal  $S(t)$ . Solving equation 5.2 for  $P$  is an ill-posed problem because the kernel matrix  $K$  is singular (or nearly so), that is its rank  $k < M$ . Furthermore,  $S(t)$  contains noise, increasing the challenge of obtaining a reliable  $P(r)$ .

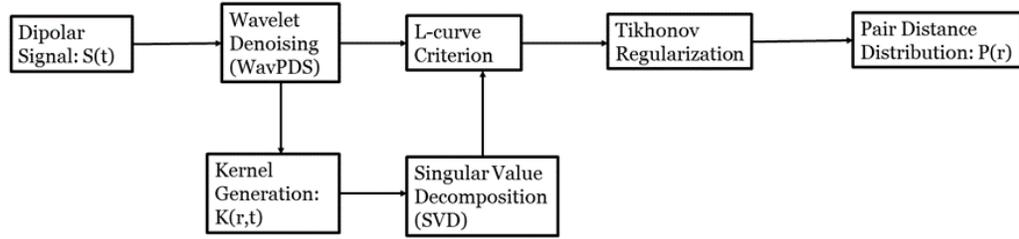
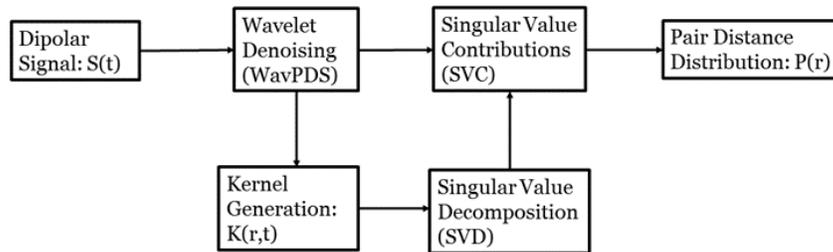
### 5.3 Current Methods for Distance Distribution Reconstruction

Thus, as commented above, one usually resorts to regularization or model fitting methods to obtain reasonable results for  $P(r)$ , and the method of Tikhonov regularization (TIKR) is commonly used.<sup>7,9</sup> But it relies heavily on the selection of the regularization parameter ( $\lambda$ ), which yields a compromise between a reliable fit to the data in a least squares sense and maintaining stability of the  $P(r)$ .<sup>7,9,24</sup> For a noisy dipolar signal, the TIKR method compromises on resolution in order to avoid an unstable  $P(r)$ . Moreover, this procedure is vulnerable to the appearance of spurious peaks and negative values for  $P(r)$ , due in part to noise in the dipolar signal.

Model fitting methods that are also used<sup>9,23,33–35,25–32</sup> require, on the other hand, *a priori* model functions to estimate  $P(r)$ , which may not accurately represent the actual distance distributions. This becomes especially true if the  $P(r)$  is multimodal. Model fitting methods, as well as TIKR, are frequently unable to differentiate between intrinsically non-overlapping peaks but which are significantly broadened by the needs of the methodology, as well as by highly “dynamical” systems characterized by a broad distance distribution when frozen.

## 5.4 SVD-based Method

We now describe our *objective* method (which can be easily automated) based on SVD to resolve such matters. It yields accurate distance distributions with high resolution and without spurious peaks or negative  $P(r)$ 's, without resorting to regularization or model fitting. The dipolar signal is first denoised using our recently developed WavPDS method<sup>36</sup> based on wavelet transforms.<sup>37</sup> (Alternatively, a signal already with high SNR, e.g. from signal averaging, could skip this first step). This denoised signal is then directly converted into the  $P(r)$  by our new SVD approach. This scheme is shown in Scheme 5.1 and compared with the method using Tikhonov regularization.

**A****B**

**Scheme 5.1:** Block Diagram showing the determination of distance distributions from pulsed dipolar spectroscopy using A) Tikhonov regularization and B) the new Singular Value Decomposition Method.

In its basic form, the SVD of the “Kernel” matrix  $K$  can be obtained as  $K = U\Sigma V^T$  so equation 5.2 can be rewritten as,<sup>7</sup>

$$U\Sigma V^T P = S \tag{5.3}$$

where  $U$  is an  $M \times M$  column-orthogonal matrix,  $V$  is an  $N \times N$  column-orthogonal matrix, and  $\Sigma$  is an  $M \times N$  diagonal matrix containing the non-negative singular values,  $\sigma_i$  (where  $i = 1, 2, \dots, N$ ) in decreasing order.<sup>38</sup>

To obtain  $P$ , using the SVD in equation 5.3 one has:

$$P = V\Sigma^{-1}U^T S \tag{5.4}$$

where  $\Sigma^{-1}$  represents the diagonal matrix containing the reciprocal of the singular values. Then equation 5.4 can be rewritten in discretized form as,

$$P_j = \sum_{i=1}^k \sum_{l=1}^M V_{ji} (\Sigma^{-1})_{ii} U^T_{il} S_l \tag{5.5}$$

where  $j$  is the index representing distance  $r$ ,  $l$  is the index representing time  $t$ , and  $i$  is the index representing the singular values. The sum is only over the  $k$  positive non-zero  $\sigma_i$ 's, where  $k$  is the rank of  $K$ , whereas the  $N - k$  remaining  $\sigma_i$ 's should be zero representing their linear dependence. Thus one can, in principle, completely

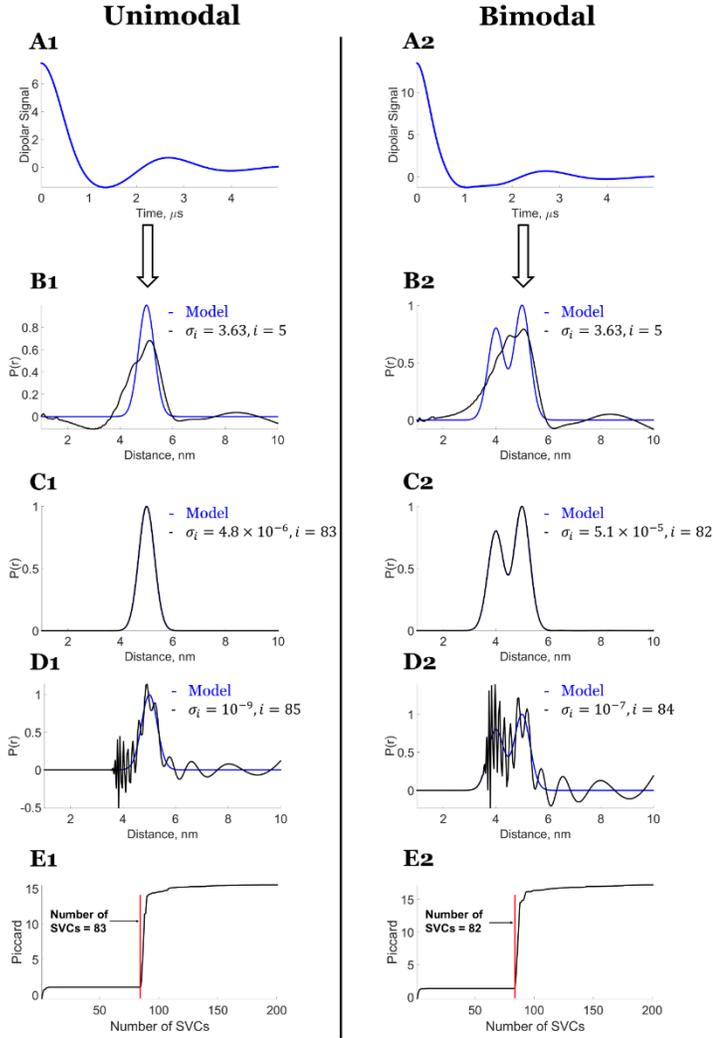
characterize the solution  $P$  by an analysis of the singular value contributions (SVCs):

$$\frac{V_{ji} U^T_{il} S_l}{\sigma_i}$$

in the sum of equation 5.5, where  $(\Sigma^{-1})_{ii} = \frac{1}{\sigma_i}$ , for the non-zero  $\sigma_i$ 's. But, in

practice, it is difficult to determine  $k$  for the ill-posed problem even with a noise-free signal, due to computational round-off noise, so that the  $(N - k)$  singular values are not precisely zero. To obtain a good solution  $P$ , a metric is needed to determine the cut-off value for the  $\sigma_i$ .

One such approach is simply to start with just the SVC from the largest singular value (SV) in equation 5.5. Then sequentially add the SVC's from successively smaller SV's while monitoring the  $i^{th}$  approximation to  $P_j$  given by equation 5.5. As one does this the  $P(r)$  becomes better approximated. We illustrate this in Fig. 5.1 for the ideal cases of noise-free  $P(r)$  models of 1) unimodal and 2) bimodal distributions. Here A1 & A2 show their dipolar signals in the time domain. The B1 & B2 compare the approximations to  $P(r)$  after just 5 SVC's to the model shown in blue. The C1 & C2 show the  $P(r)$  when they have converged to the exact result for  $i = 83$  and 82 respectively, whereas D1 and D2 show the divergent results for  $P(r)$  with just two additional SVC's. Thus one method is to add SVC's sequentially and to stop just before additional SVC's lead to diverging results. This is shown in more detail in the SI Figs. 5.S4, and 5.S6. In fact these results are fully consistent with the theorem<sup>12</sup> stating the necessary and sufficient conditions for obtaining the exact results by SVD, that is given in the SI and illustrated in Figs. 5.S1 and 5.S2.



**Figure 5.1:** Model Data -- The exact solution using the SVD method;  $P(r)$  vs. Number of Singular Value Contributions (SVCs). Unimodal Distance Distribution: **A1)** Noise-free dipolar signal. Comparison of model distribution with  $P(r)$  generated from: **B1)** fewer SVCs  $i = 3$ ,  $\sigma_i = 8.04$ ; **C1)** exact solution obtained for  $i = 83$ ,  $\sigma_i = 4.8 \times 10^{-6}$ ; **D1)** more SVCs  $i = 85$ ,  $\sigma_i = 10^{-9}$ ; **E1)** Piccard plot of  $\log_{10} \left( \sum_i |u_i^T S / \sigma_i|^2 \right)$  vs. the number of singular values from  $i = 1$  to 200 starting from largest value; it shows the contributions of singular values that lead to stable and unstable distributions. Bimodal Distance Distribution: **A2)** Noise-free Dipolar Signal. Comparison of model distribution with  $P(r)$  generated from: **B2)** fewer SVCs  $i = 3$ ,  $\sigma_i = 8.04$ ; **C2)** exact solution obtained for  $i = 82$ ,  $\sigma_i = 5.1 \times 10^{-5}$ ; **D2)** more SVCs  $i = 84$ ,  $\sigma_i = 10^{-7}$ ; **E2)** Piccard plot of  $\log_{10} \left( \sum_i |u_i^T S / \sigma_i|^2 \right)$  vs. the number of singular values from  $i = 1$  to 200 starting from largest value; it shows the contributions of singular values that lead to stable and unstable distributions.

There is a second approach to determining the cut-off value for the SVC's. This is based on the Piccard condition,<sup>7,39</sup> which requires that in order for equation 5.5 to have a solution, it is necessary and sufficient that

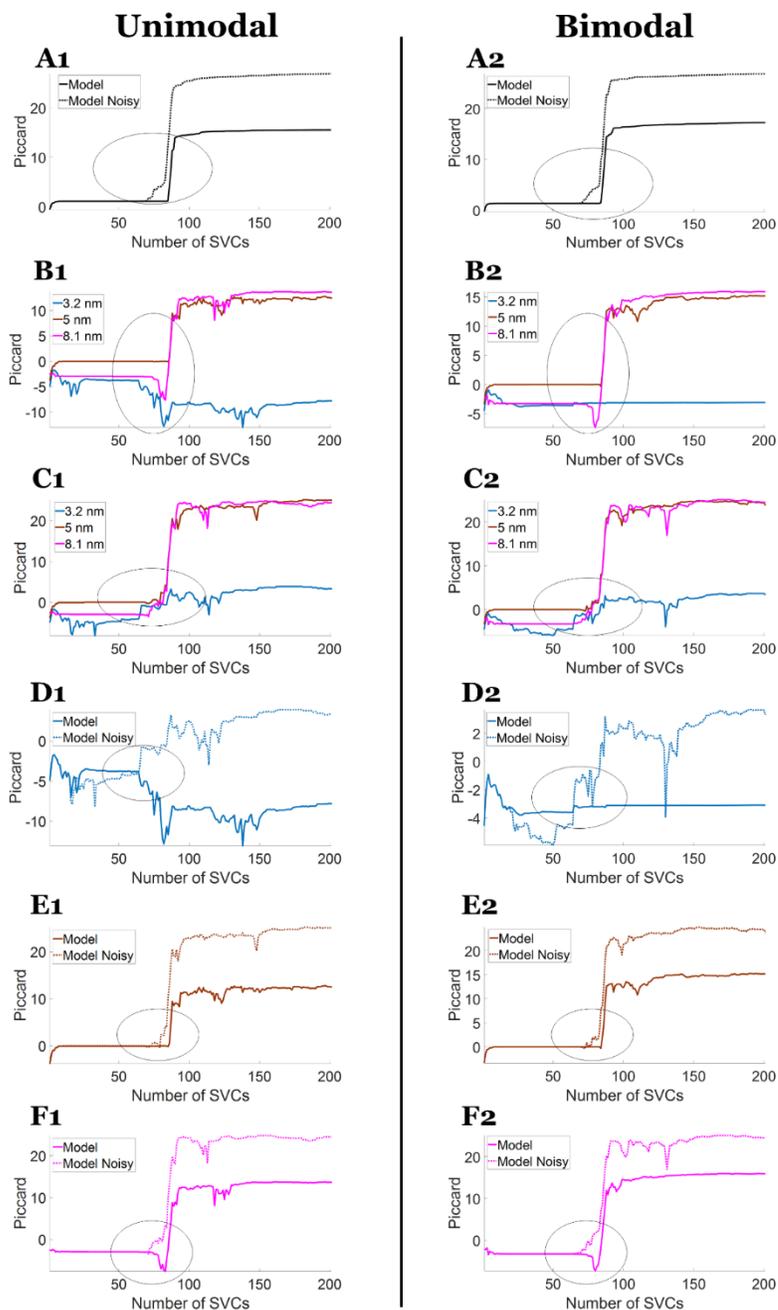
$$[Piccard] \equiv \sum_{i=1}^q \frac{1}{\sigma_i^2} |\langle S, u_i \rangle|^2 < \infty \quad (5.6)$$

where  $q \leq k < M$ , and  $\langle S, u_i \rangle$  is the inner product of the two vectors and  $u_i$  is the  $i^{th}$  column of the column-orthogonal matrix  $U$ . The Piccard condition states that to obtain a desirable solution, equation 5.6 should converge to a finite value, which is represented by less than  $\infty$  sign. It also informs one when a particular solution begins to diverge with the addition of further singular values. For the noise-free examples we are first addressing, we can expect  $q = k$  is satisfactory but  $q > k$  is not. This is shown in Fig. 5.1, E1 & E2 which show that the (log of the) Piccard sum diverges at the number of SVC's exactly where the  $P(r)$  approximations become distorted. Thus one can use the cut-off value of  $i$  as that just before the Piccard plot diverges.

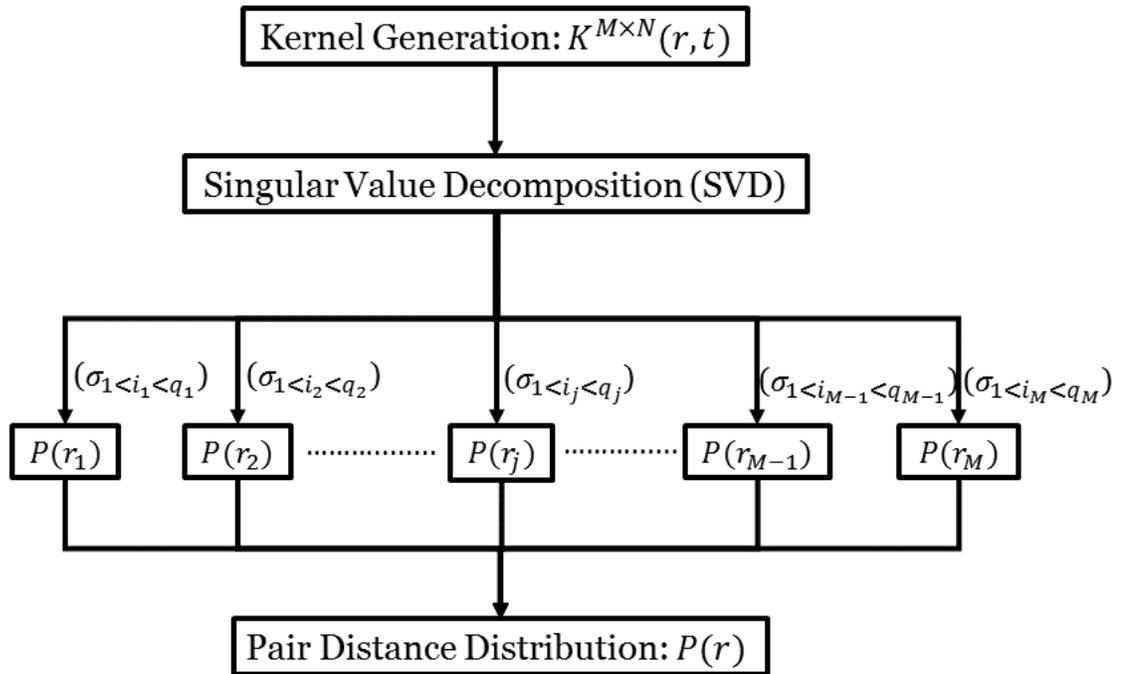
If only a small amount of noise is added (we use a nominal value of  $SNR = 850$ , which is consistent with our denoising capabilities<sup>36</sup>), then the necessary and sufficient conditions of the theorem are no longer satisfied, as also illustrated in Figs. 5.S1 and 5.S2; so the SVD can only be used to obtain a best fit in a least squares sense.<sup>12</sup> Also, in the presence of even such modest noise not only must one cut off the sum of SVC's in equation 5.5 sooner, but additionally there is no longer a precise cut-off for the

SVC's as shown in the Piccard plots of Figs. 5.2 A1 and A2; there is a blurring of this cut-off as is illustrated further in plots of  $P(r)$  vs. added SVC's shown in Figs. 5.S5 and 5.S7 for  $SNR = 850$ .

We have found that this is due to the fact that a good choice of cut-off becomes dependent upon the value of  $r$ , and this is illustrated in Fig. 5.2 for Piccard plots done separately for  $r = 3.2, 5, \text{ and } 8.1 \text{ nm}$ . For a noise-free signal, in Fig. 5.2 B1 and B2, one can see that different values of  $r$  diverge at about the same singular value cut-off or have already converged and remain so, (e.g.  $r = 3.2$ ). However, there is a change once noise is present. Then each  $r$  value diverges at different singular value cut-offs (cf. Fig. 5.2 C1 and C2) in addition to the fact that the cut-offs occur sooner as a function of  $i$  than in the noise-free case (cf. Fig. 5.2 E1, E2, F1 and F2). In Fig. 5.2 D1 and D2 at  $3.2 \text{ nm}$ , the Piccard plot for a noisy dipolar signal diverges at  $\sigma_i = 0.2$ , whereas for the noise-free signal it remains converged. It can also be seen from Fig. 5.2, that shorter distances diverge sooner than longer distances. Scheme 5.2 shows the block diagram of the SVD based approach using different SVCs for different  $r_j$  values, such that  $q \rightarrow q_j$ . If this fact is ignored, spurious peaks and other defects will arise as illustrated in the SI (Figs. 5.S9-5.S12 and below).



**Figure 5.2:** Piccard plots of  $\log_{10} \left( \sum_i |u_i^T S / \sigma_i|^2 \right)$  comparing noise-free and some noise ( $SNR = 850$ ) cases for different distance segments. Case 1 is Unimodal and Case 2 is Bimodal Model. **A)** Piccard plots for noise-free vs some noise. **B)** Comparison of distance ( $r$ ) dependent Piccard plots at  $3.2 \text{ nm}$ ,  $5 \text{ nm}$ , and  $8.1 \text{ nm}$  for noise-free models. **C)** Same as **B)** except for models with some noise. Comparison of distance ( $r$ ) dependent Piccard plots for noise-free vs. some noise cases for **D)**  $3.2 \text{ nm}$ ; **E)**  $5 \text{ nm}$ ; **F)**  $8.1 \text{ nm}$ .

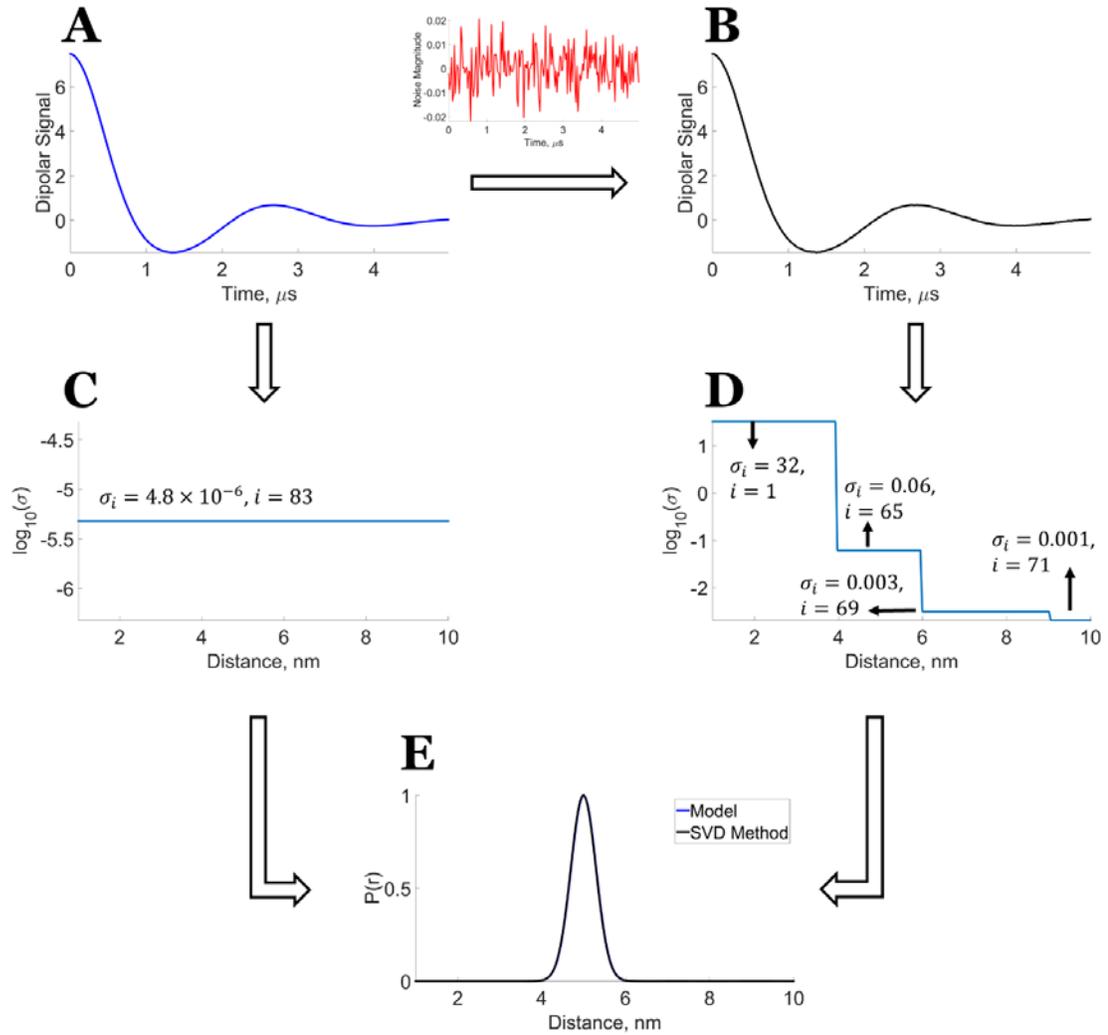


**Scheme 5.2:** Block Diagram showing generation of probability values at each distance measurement. For practical analyses, it is sufficient to consider just 3 or 4 ranges, cf. Figs. 5.3 and 5.S8.

Therefore, in principle, one needs to construct the  $P_j$  for each  $r_j$  value separately using equation 5.5. Equation 5.5 can thus be rewritten to represent this approach as,

$$P_j = \sum_{l=1}^M \sum_{i=1}^{q(j)} V_{ji} (\Sigma^{-1})_{ii} U^T_{il} S_l \quad (5.7)$$

However, more practically one can use just 3 or 4 segments of  $r$  as illustrated in Figs. 5.3D and 5.S8. Within each of these segments the Piccard expression diverges just after the correct  $q_j$  is reached as shown in Figs. 5.S13-5.S20. In a few cases, some  $r$  values in a segment do not yet diverge at this cut-off (Figs. 5.S14 and 5.S18), but they have reached their converged value such that the contribution of the subsequent  $q + 1^{th}$  singular value is negligible. The singular value cut-off is usually taken at the first  $r$  value of the segment that diverges. Remaining  $r$  values, although not yet diverged, have typically achieved their convergence (see flat lines in Figs. 5.S14 and 5.S18). It is this property that enables the segmentation approach. Of course, one can further divide that segment into two or more segments and select their respective singular value cut-offs, but we find it is unnecessary. In some cases, especially at short distances for which  $P(r) = 0$  (e.g. Figs. 5.S13 and 5.S17), divergence occurs after only a very few SVCs. We have noticed that the presence of noise in the dipolar signal severely affects the short distances where  $P(r) = 0$ . One can see in the insets to Figs. 5.S13 and 5.S17 that the Piccard plot in the segment diverges from just the first SVC to the second SVC. In Fig. 5.S13, the Piccard plots further diverge at later SVC. This subsequent divergence clearly is not important.



**Figure 5.3:** Reconstruction of distance distribution for noise-free model data and noisy model data ( $SNR \approx 850$ ) using the new SVD method. **A)** Model dipolar signal; **B)** Model dipolar signal with added noise (see added noise in Red plot); **C)** Singular value cut-off at each distance ( $nm$ ) for the model dipolar signal; **D)** Singular value cut-off at each distance ( $nm$ ) for the model dipolar signal with added noise; and **E)** Distance distribution reconstructed from the model dipolar signal and model dipolar signal with noise using the singular value cut-offs shown in **C)** and **D)**, respectively. Note that the added noise is so small that **A** and **B** still appear identical, but convergence to the virtually identical final results requires segmentation in the latter case.

To obtain appropriate segmentation and singular value cut-off for each segment, one first performs the sum of equation 5.5 starting from the largest singular value with the whole distribution taken as a single segment. SVCs are sequentially added in the order of decreasing singular values until the  $i^{th}$  SVC yields an unstable approximation to  $P(r)$  for a given region of  $r$ . That region of  $r$  becomes the first segment with no further SVCs needed. SVCs are then added for the remaining  $r$  region(s) until another region of  $r$  become unstable. This process is repeated until convergence for all the regions of  $r$  is achieved. This SVC approach leads to a good approximation of  $P(r)$  without spurious peaks.

Alternatively, one can study the Piccard plot using equation 5.6 to identify segments and to obtain singular value cut-offs for each segment. In the Piccard plot, segments represented by the regions of  $r$  will collectively diverge at different SVCs, informing about the number of segments needed. A singular value cut-off for each of the segments can then be independently calculated by adding the SVC's for the  $j^{th}$  segment to locate the onset of divergences occurring at  $q_j + 1$ , thereby giving the correct  $q_j$  for each segment. This approach is the more convenient and lends itself to automation in programming.

## 5.5 Results and Discussion

### 5.5.1 Experimental and Model Data

We tested the new SVD method on model data as well as experimental data from ref. 14 and compared it with the TIKR method as well as with TIKR followed by the maximum entropy method (MEM)<sup>8</sup> (see Figs. 5.S9-5.S12). The MEM method improves the fit and removes the negative values of  $P(r)$ . For the model cases, we used both a unimodal and bimodal Gaussian distribution with peaks respectively at 4 nm, and 4 nm and 5 nm, each having a standard deviation of 0.3 nm (see Figs. 5.S9-5.S10). For the bimodal case, the height of the first peak is taken as 80% of the second peak. The evolution time for both unimodal and bimodal dipolar signal is 5  $\mu$ s. The initial SNR of 30 and 10 respectively was generated using white Gaussian noise, but we just show the denoised version.

For experimental cases, we used unimodal and bimodal distributions (Fig. 5.S11 and Fig. 5.S12, respectively) generated from a T4 lysozyme (T4L) system. The sample for the unimodal distribution was spin-labeled with T4L mutant 44C/135C at 63  $\mu$ M concentration. Signal was acquired for 112 min and 14 min to obtain different SNRs. The sample for the bimodal distribution was spin-labeled with T4L mutant 8C/44C and 44C/135C at 44  $\mu$ M and 47  $\mu$ M concentration, respectively. For different SNRs, signal averaging of 48 min and 8 min was carried out. The dipolar signal was obtained from Double Electron-Electron Resonance (DEER) experiments at 17.3 GHz. For

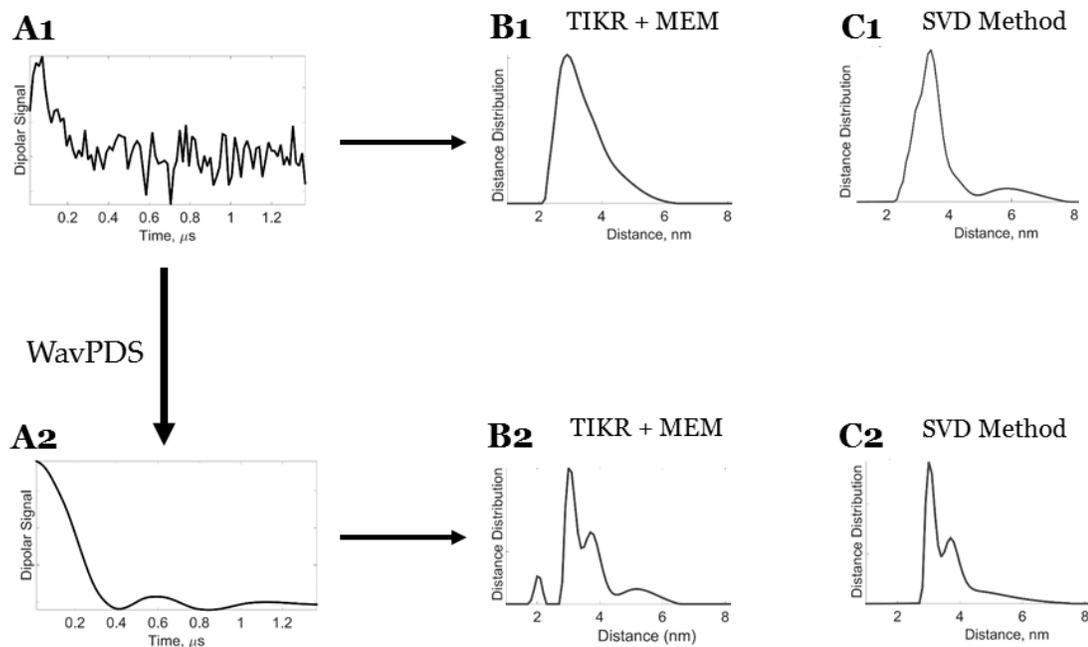
further details about the experiment, see ref. 14. Figs. 5.S9 to 5.S12 show the successful application of the SVD method.

### **5.5.2 WavPDS**

The WavPDS method was used to first denoise the model and experimental dipolar signals to obtain high SNR (see Table 5.S1). One sees in this Table that the SNR's of all the examples we used range from 488 to 3,333. In all these cases, despite the very large SNR's, the segmentation method described above was necessary.

### **5.5.3 Experimental Case Study**

In Fig. 5.4 we illustrate a challenging experimental case, where before denoising the  $SNR = 3.8$ , whereas after denoising it becomes 488. We also show in this figure (as well as in Figs. 5.S9-5.S12) that the standard TIKR+MEM of the denoised signal yields to spurious peaks, which however are non-existent with the SVD method proposed here. Spurious peaks at short distances occur because the distribution has diverged (become unstable) at those  $r$  values, whereas spurious peaks at long distances occur because the distribution has not converged yet at those  $r$  values. The TIKR and TIKR+MEM methods have a single singular value cut-off for all  $r$  values, which results in these spurious peaks. The new SVD method overcomes the problem by selecting different singular value cut-offs (Piccard Plots: Figs. 5.S21 and 5.S22). Also, shown in Fig. 5.4 is the fact that before denoising neither the TIKR+MEM, nor the new SVD method yield satisfactory results.



**Figure 5.4:** Reconstruction of distance distribution using Tikhonov regularization + Maximum entropy method (TIKR+MEM) and new SVD method for noisy and denoised (WavPDS) experimental dipolar signal. **A1)** Noisy experimental dipolar signal from spin labeled IgE cross-linked with DNA-DNP ligand in PBS buffer solution. The dipolar signal was collected after 18 hours of signal averaging and has  $SNR = 3.8$ ; **B1)** Distance distribution obtained from noisy dipolar signal using TIKR+MEM; **C1)** Distance distribution obtained from noisy dipolar signal using new SVD method; **A2)** Denoised experimental dipolar signal using WavPDS; **B2)** Distance distribution obtained from denoised dipolar signal using TIKR+MEM; **C2)** Distance distribution obtained from denoised dipolar signal using new SVD method.

## 5.6 Conclusion

In summary, convergence of a SVD solution requires one to include all the “relevant” SVC’s, and to exclude those that are inappropriate, and this must be performed in an  $r$ -dependent fashion to obtain optimum  $P(r)$ ’s. We have described such a method in this letter.

## 5.S Supplementary Information

### 5.S.1 Details of Platform and Software

A machine with Intel Core i5-4590 CPU @ 3.30 GHz processor, Windows 7 operating system, 16 GB RAM, and 64-bit operating system was used as the platform. Denoising was performed using MATLAB 2014b. Tikhonov Regularization (TIKR)<sup>7</sup> and Maximum Entropy Method (MEM)<sup>8</sup> codes written for MATLAB available on ACERT website ([https://acert.cornell.edu/index\\_files/acert\\_resources.php](https://acert.cornell.edu/index_files/acert_resources.php)) were used to generate  $P(r)$ .

### 5.S.2 Theorem for Obtaining Exact Solution Using SVD

For a system of linear equations  $KP = S$ , where  $K$  is an  $M \times N$  matrix ( $N \leq M$ ) with a rank  $k$  ( $k < M$ ), where  $P$  is a vector with length  $N$ , and  $S$  is a vector with length  $M$ , then there exists an exact solution using singular value decomposition (SVD) if and only if  $S$  is orthogonal to the  $M - k$  left-singular vectors of  $K$ .<sup>12</sup> To illustrate, in SVD form,  $KP = S$  can be written as (cf. eq. 5.3),

$$U\Sigma V^T P = S \tag{5.S1}$$

where  $U$  is an  $M \times M$  column-orthogonal matrix,  $V$  is an  $N \times N$  column-orthogonal matrix, and  $\Sigma$  is an  $M \times N$  diagonal matrix containing the non-negative singular values ( $\sigma$ ) in decreasing order. To solve for  $P$ , equation S1 can be rewritten as (cf. eq. 5.4),

$$P = V\Sigma^{-1}U^T S \tag{5.S2}$$

or,

$$P = \sum_{i=1}^N \frac{u_i^T S}{\sigma_i} v_i \quad (5.S3)$$

or,

$$P = \sum_{i=1}^k \frac{u_i^T S}{\sigma_i} v_i + \sum_{i=k+1}^N \frac{u_i^T S}{\sigma_i} v_i \quad (5.S4)$$

where  $u_i$  and  $v_i$  are the column vectors of the matrices  $U$  and  $V$ , respectively,  $\sigma_i$  are the individual singular values of the diagonal matrix  $\Sigma$ .

If  $u_i^T S = 0, \forall i \in [k + 1, m - k]$ , then the linear system has a solution for  $P$  and can be obtained as,

$$P = \sum_{i=1}^k \frac{u_i^T S}{\sigma_i} v_i \quad (5.S5)$$

Pages 428-429 in ref 4 provide a more detail description.

### 5.S.3 Tikhonov Regularization, its SVD form and its Drawback

Tikhonov regularization (TIKR) minimizes the following function,<sup>7,8</sup>

$$\phi_{TIKR}[P] \equiv \min \|KP - S\|^2 + \lambda^2 \|LP\|^2 \quad (5.S6)$$

where  $\lambda$  is the regularization parameter and  $L$  is a differential operator. The solution  $P_\lambda$  can be obtained as,

$$P_\lambda = (K^T K + \lambda^2)^{-1} K^T S \quad (5.S7)$$

Equation 5.S7 can be reduced to SVD form in the following way,

$$P_\lambda = (\Sigma^2 + \lambda^2 I)^{-1} V \Sigma U^T S \quad (5.S8)$$

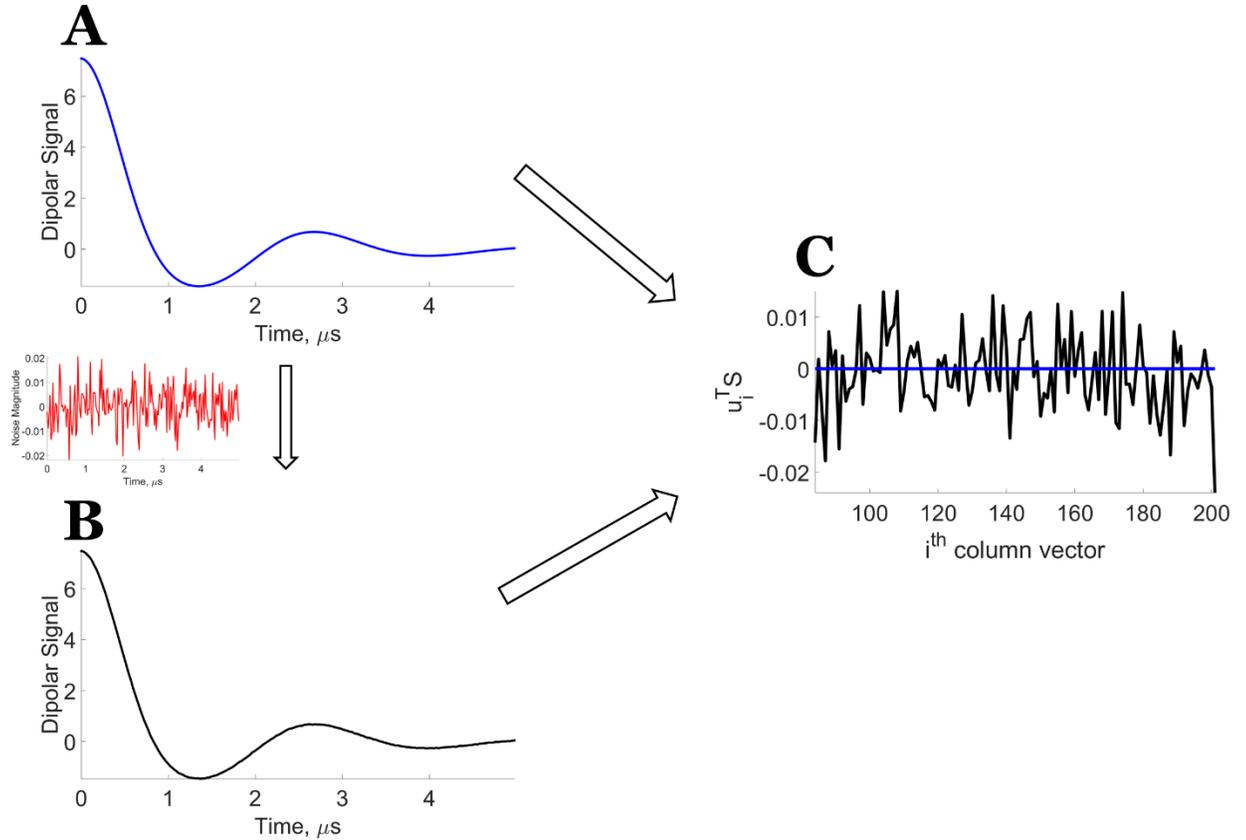
$P_\lambda$  in equation 5.S8 can be rewritten as a modified sum of singular value contributions (SVCs) as,

$$P_\lambda = \sum_{i=1}^{\sigma_i \geq \lambda} f_i \frac{u_i^T S}{\sigma_i} v_i \quad (5.S9)$$

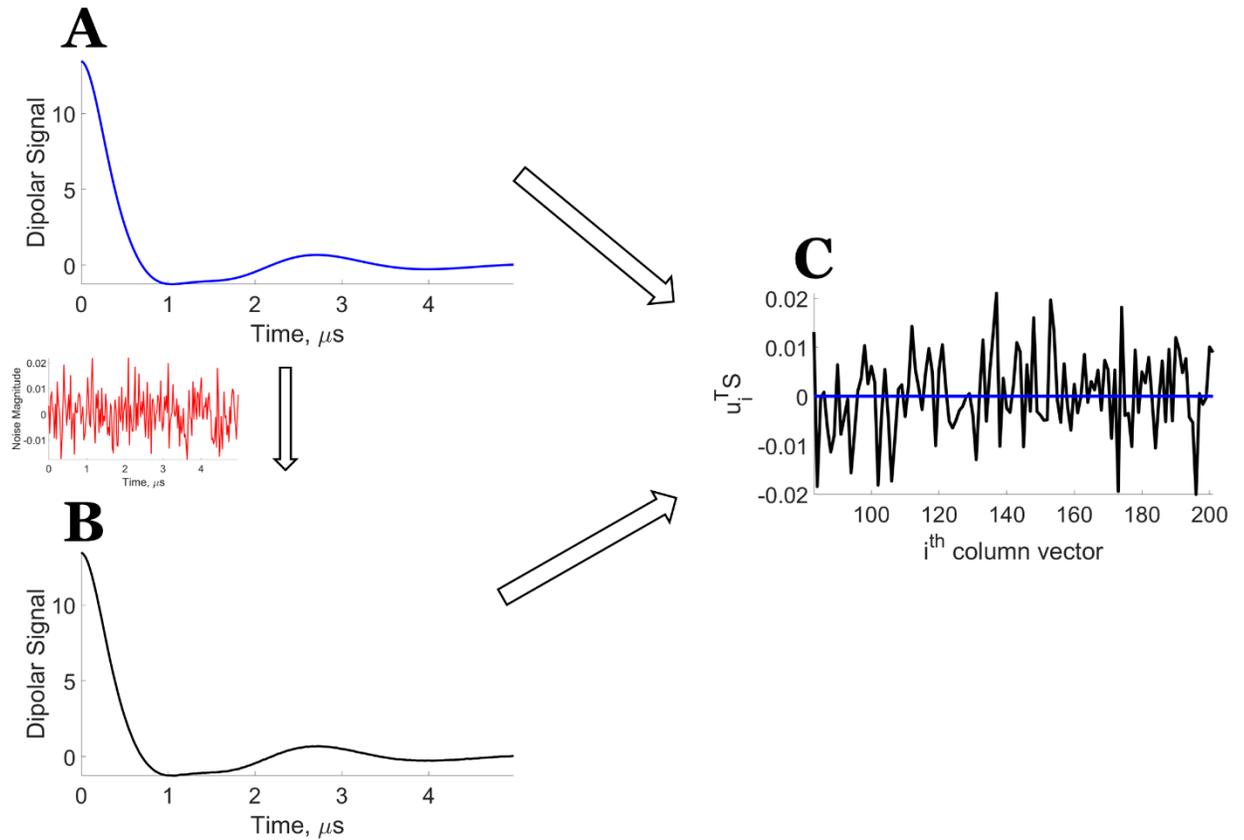
where  $f_i \equiv \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2}$  and is a filter function that suppresses the SVCs from small singular values such that  $\sigma_i < \lambda$ . The choice of  $\lambda$  is crucial; one such approach is to obtain it by the  $L$ -curve method.<sup>24,40</sup> Scheme 5.1A shows a block diagram of the TIKR approach. The use of  $f_i$  provides a “softer” suppression than the hard suppression in the new SVD method where the  $q = i^{th}$  SVC is kept and all SVC’s for  $q > i$  are set to zero.

TIKR cannot be implemented with  $r$  dependence, because it solves a least square problem using a single  $\lambda$ . The method is designed to solve least squares involving the whole distribution.

**S4. Orthogonality Validation based on Theorem for Unimodal and Bimodal Model Data: Figs. 5.S1, 5.S2**

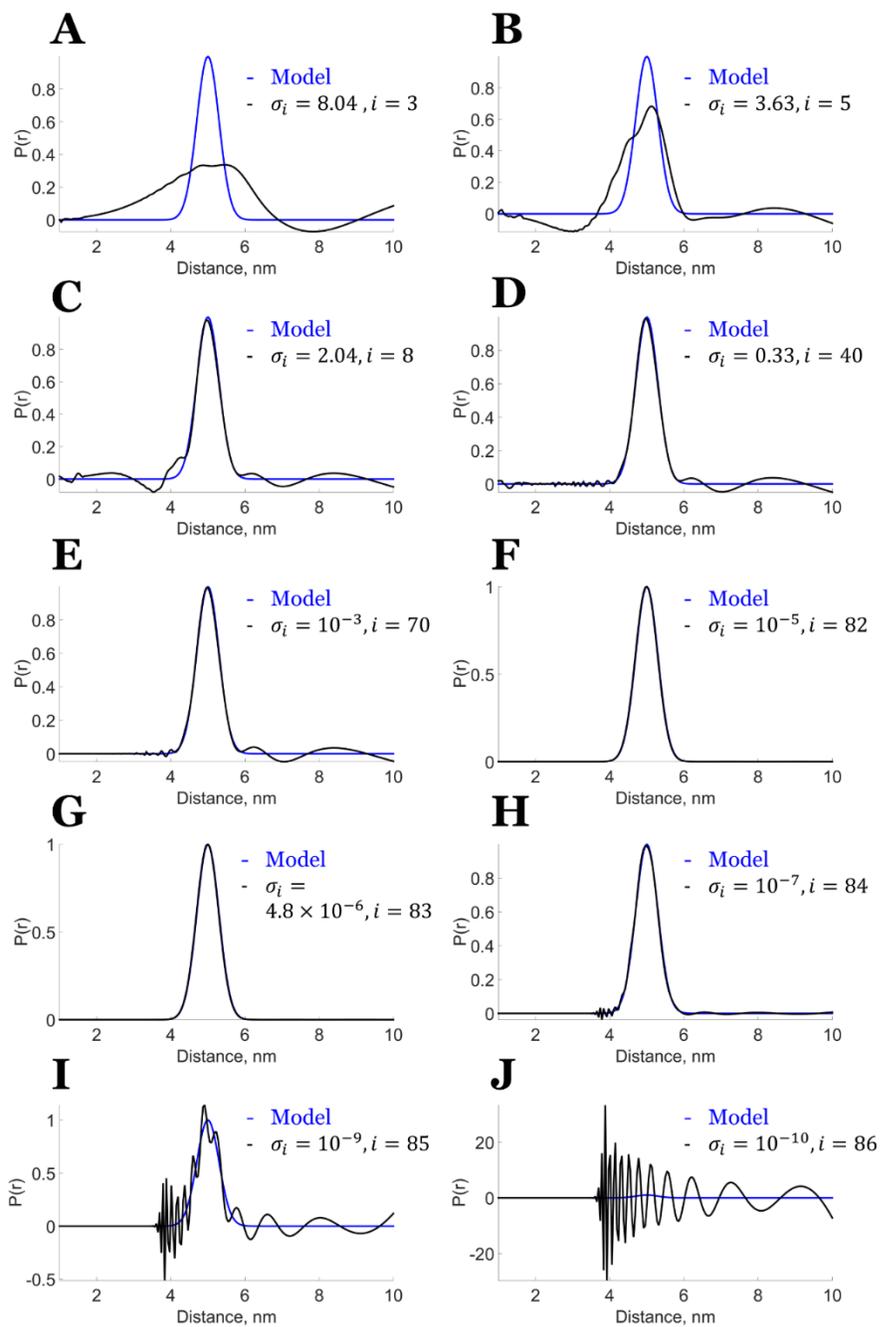


**Figure 5.S1:** Unimodal Model Data -- Orthogonality validation for the noiseless data and data with some noise ( $SNR \approx 850$ ). **A)** Noiseless data; **B)** Data with some noise; and **C)** Comparison of results of  $u_i^T S$  at  $i^{\text{th}}$  column vector for noiseless data and data with some noise. Signal in red is noise added to the noise-free data (blue) to obtain data with some noise (black). The orthogonality requirement shown in **C)** is clearly met for the noiseless data but not for data with some noise.

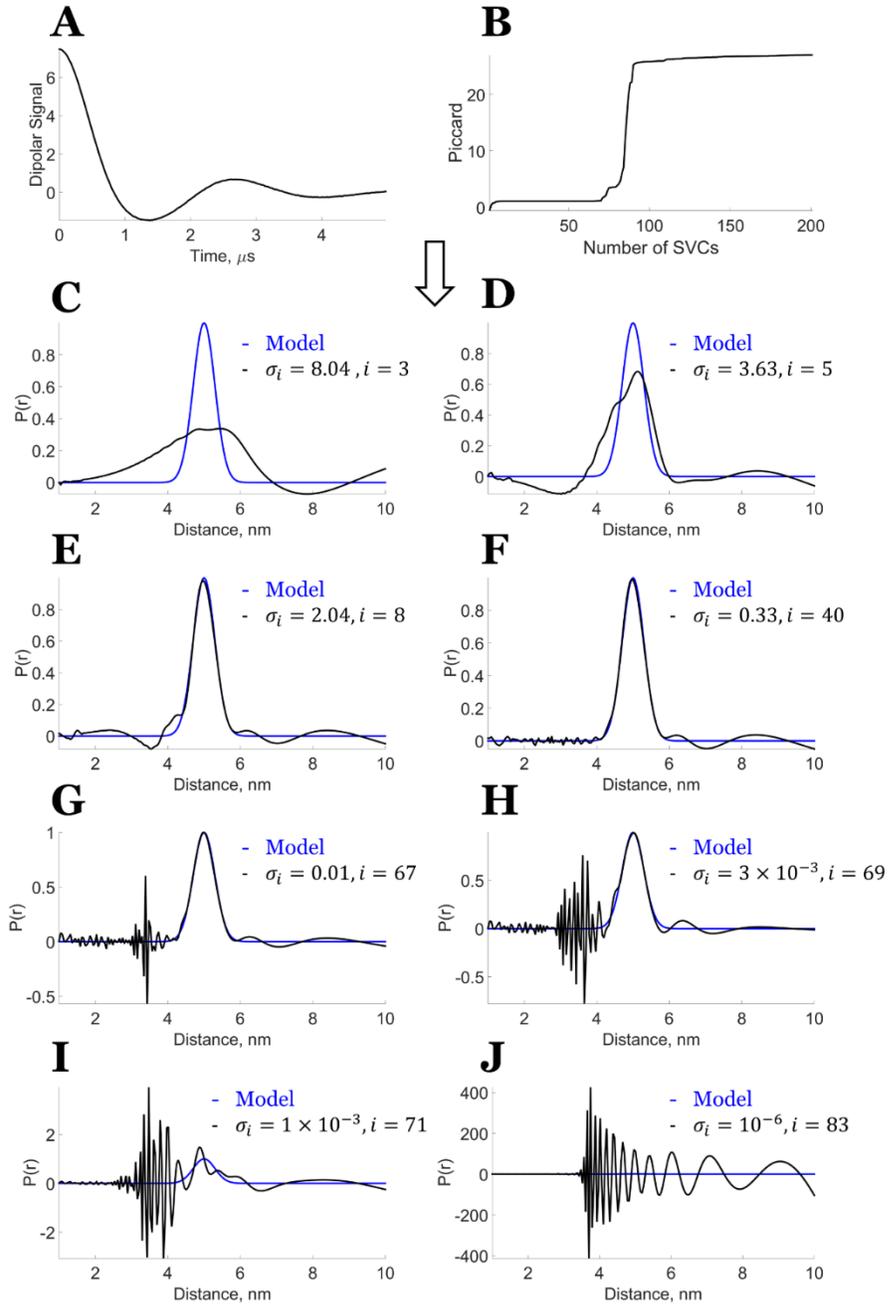


**Figure 5.S2:** Bimodal Model Data -- Orthogonality validation for the noiseless data and data with some noise ( $SNR \approx 850$ ). **A)** Noiseless data; **B)** Data with some noise; and **C)** Comparison of results of  $u_i^T S$  at  $i^{\text{th}}$  column vector for noiseless data and data with some noise. Signal in red is noise added to the noise-free data (blue) to obtain data with some noise (black). The orthogonality requirement shown in **C)** is clearly met for the noiseless data but not for data with some noise.

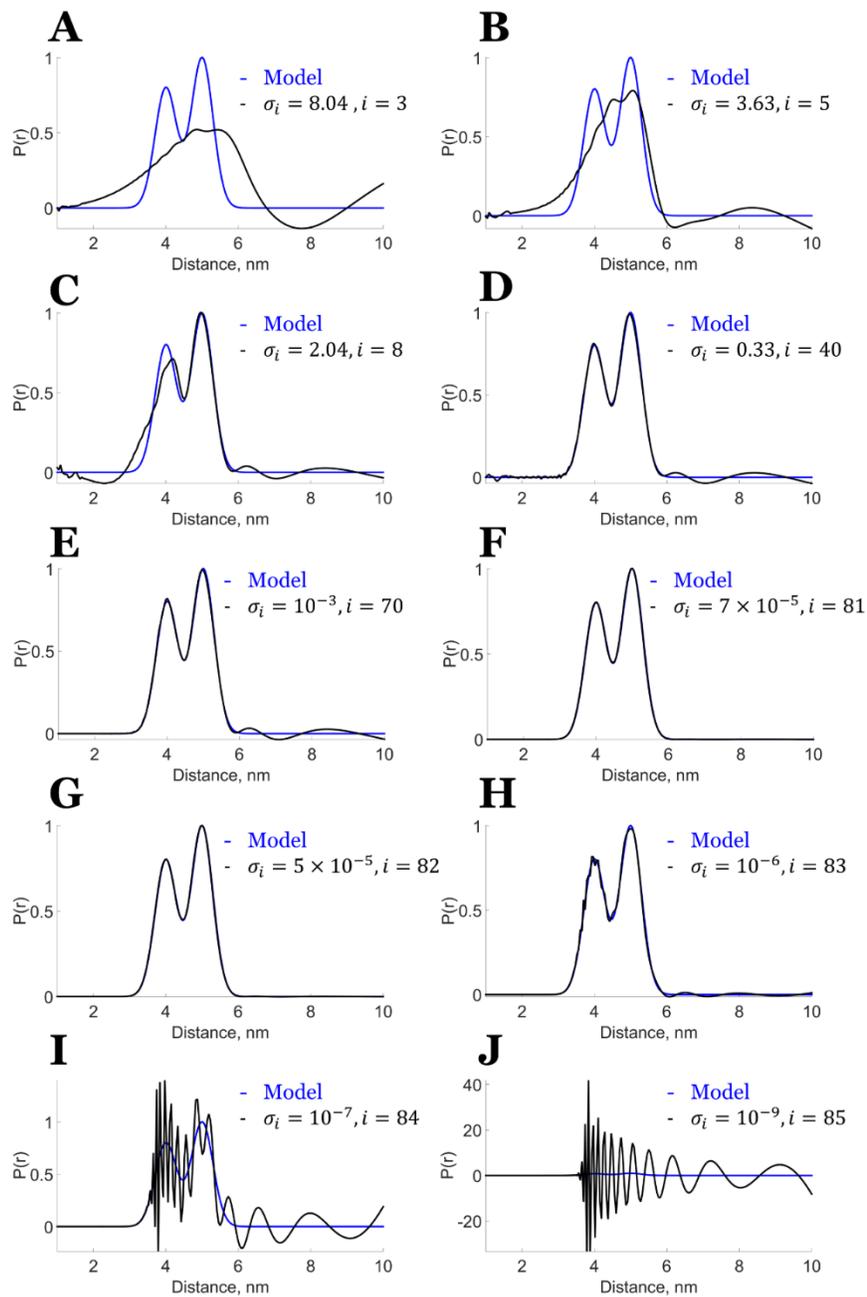
**5.S.5 Reconstructed Unimodal and Bimodal Distance Distributions of Noiseless and Some Noise Model Data with Different Singular Value Cutoffs: Figs. 5.S4-5.S7**



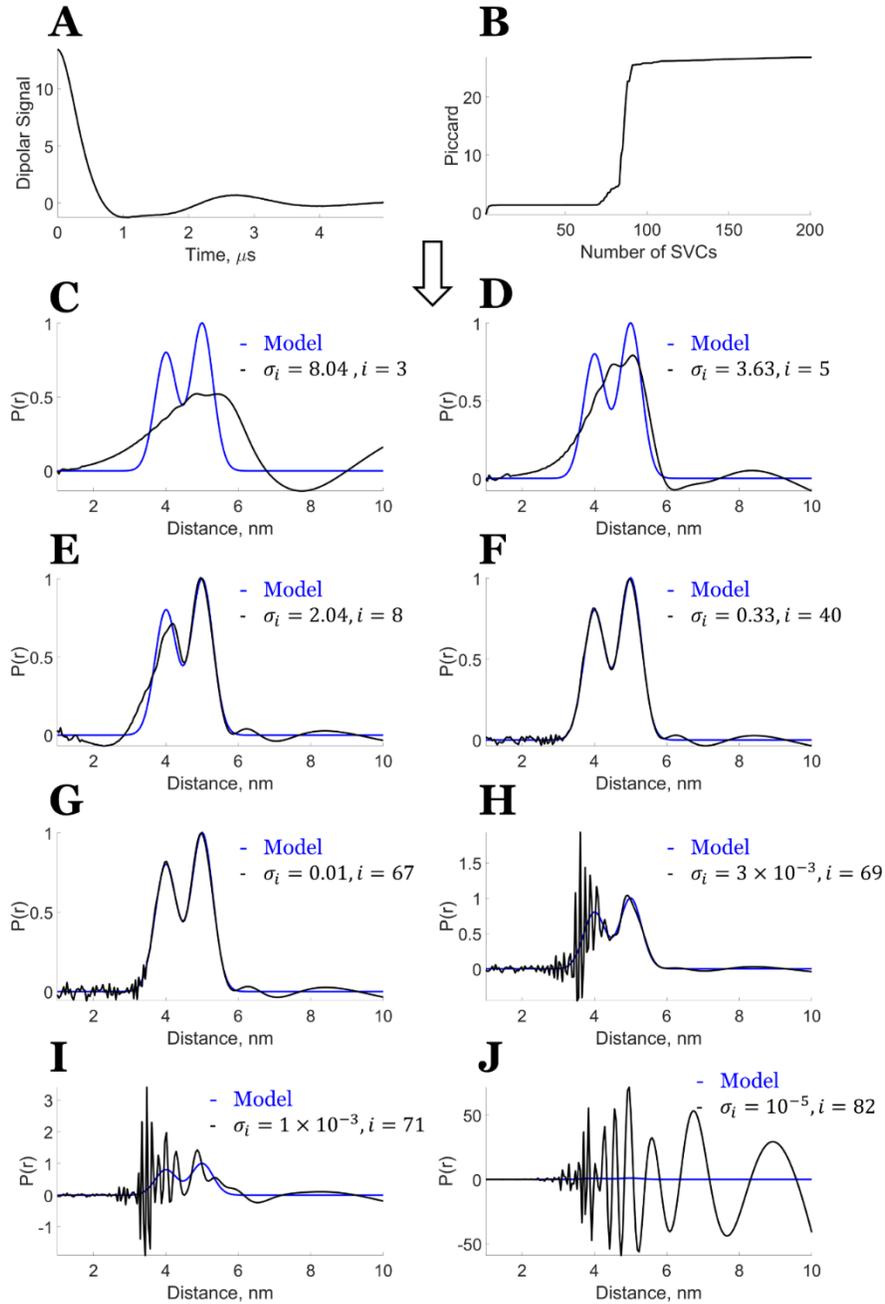
**Figure 5.S4:** Model data for noise-free case -- Unimodal Distance Distribution. Examples of  $P(r)$  obtained for different values of the singular value cut-off. **A)**  $i = 3$ ,  $\sigma_i = 8.04$ ; **B)**  $i = 5$ ,  $\sigma_i = 3.63$ ; **C)**  $i = 8$ ,  $\sigma_i = 2.04$ ; **D)**  $i = 40$ ,  $\sigma_i = 0.33$ ; **E)**  $i = 70$ ,  $\sigma_i = 10^{-3}$ ; **F)**  $i = 82$ ,  $\sigma_i = 10^{-5}$ ; **G)**  $i = 83$ ,  $\sigma_i = 10^{-6}$ ; **H)**  $i = 84$ ,  $\sigma_i = 10^{-7}$ ; **I)**  $i = 85$ ,  $\sigma_i = 10^{-9}$ ; and **J)**  $i = 86$ ,  $\sigma_i = 10^{-10}$ .



**Figure 5.S5:** Model Data with Some Noise ( $SNR \approx 850$ ) -- Unimodal Distance Distribution. **A)** Noisy model data dipolar signal; **B)** Piccard plot of the unimodal distribution from the noisy model data at different number of Singular Value Contributions (SVCs) represented by  $i$ ; and Comparison of model distribution with the distance distribution generated from: **C)**  $i = 3$ ,  $\sigma_i = 8.04$ ; **D)**  $i = 5$ ,  $\sigma_i = 3.63$ ; **E)**  $i = 8$ ,  $\sigma_i = 2.04$ ; **F)**  $i = 40$ ,  $\sigma_i = 0.33$ ; **G)**  $i = 67$ ,  $\sigma_i = 0.01$ ; **H)**  $i = 69$ ,  $\sigma_i = 3 \times 10^{-3}$ ; **I)**  $i = 71$ ,  $\sigma_i = 1 \times 10^{-3}$ ; and **J)**  $i = 83$ ,  $\sigma_i = 10^{-6}$ .

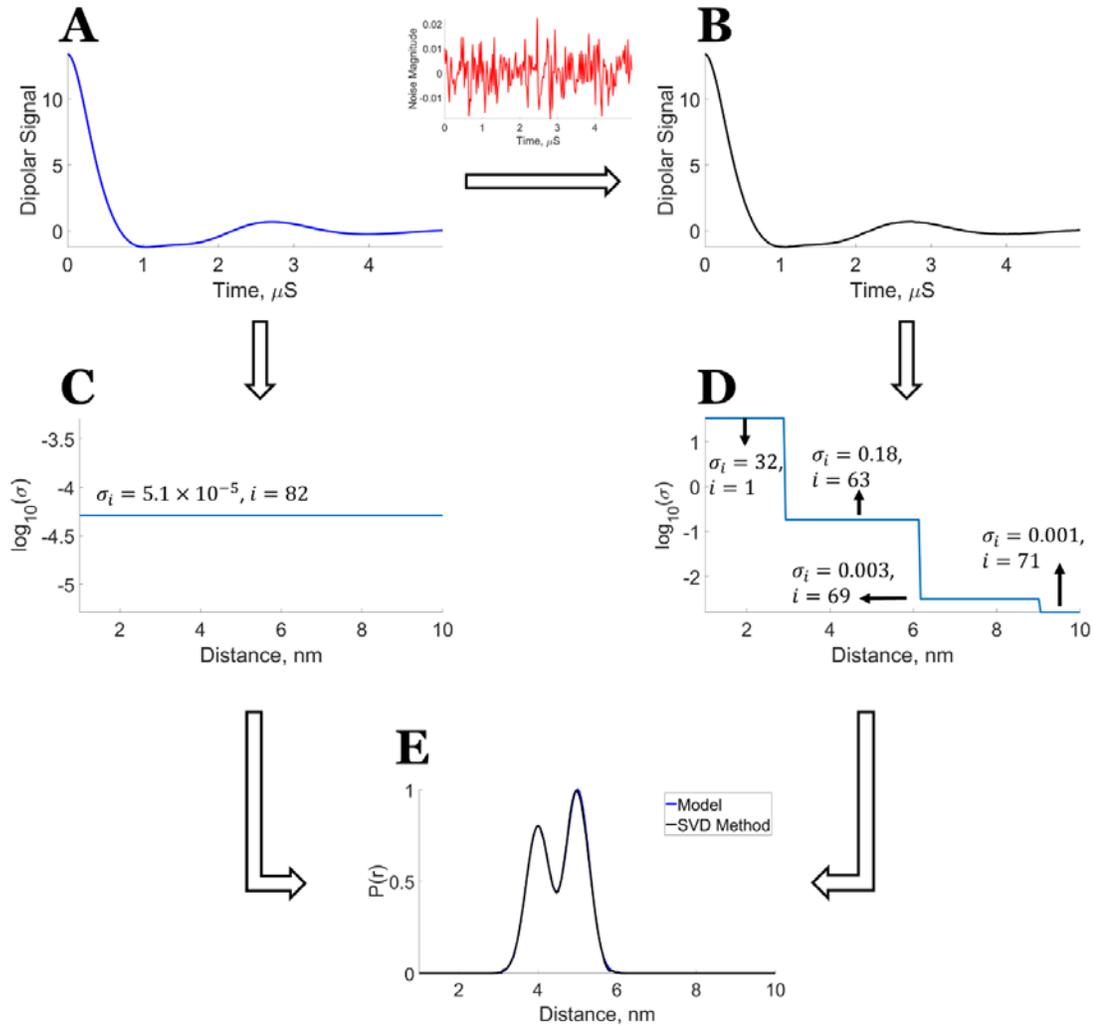


**Figure 5.S6:** Model data noise-free case -- Bimodal Distance Distribution. Examples of  $P(r)$  obtained for different values of the singular value cut-off. **A)**  $i = 3$ ,  $\sigma_i = 8.04$ ; **B)**  $i = 5$ ,  $\sigma_i = 3.63$ ; **C)**  $i = 8$ ,  $\sigma_i = 2.04$ ; **D)**  $i = 40$ ,  $\sigma_i = 0.33$ ; **E)**  $i = 70$ ,  $\sigma_i = 10^{-3}$ ; **F)**  $i = 81$ ,  $\sigma_i = 7 \times 10^{-5}$ ; **G)**  $i = 82$ ,  $\sigma_i = 5 \times 10^{-5}$ ; **H)**  $i = 83$ ,  $\sigma_i = 10^{-6}$ ; **I)**  $i = 84$ ,  $\sigma_i = 10^{-7}$ ; and **J)**  $i = 85$ ,  $\sigma_i = 10^{-9}$ .



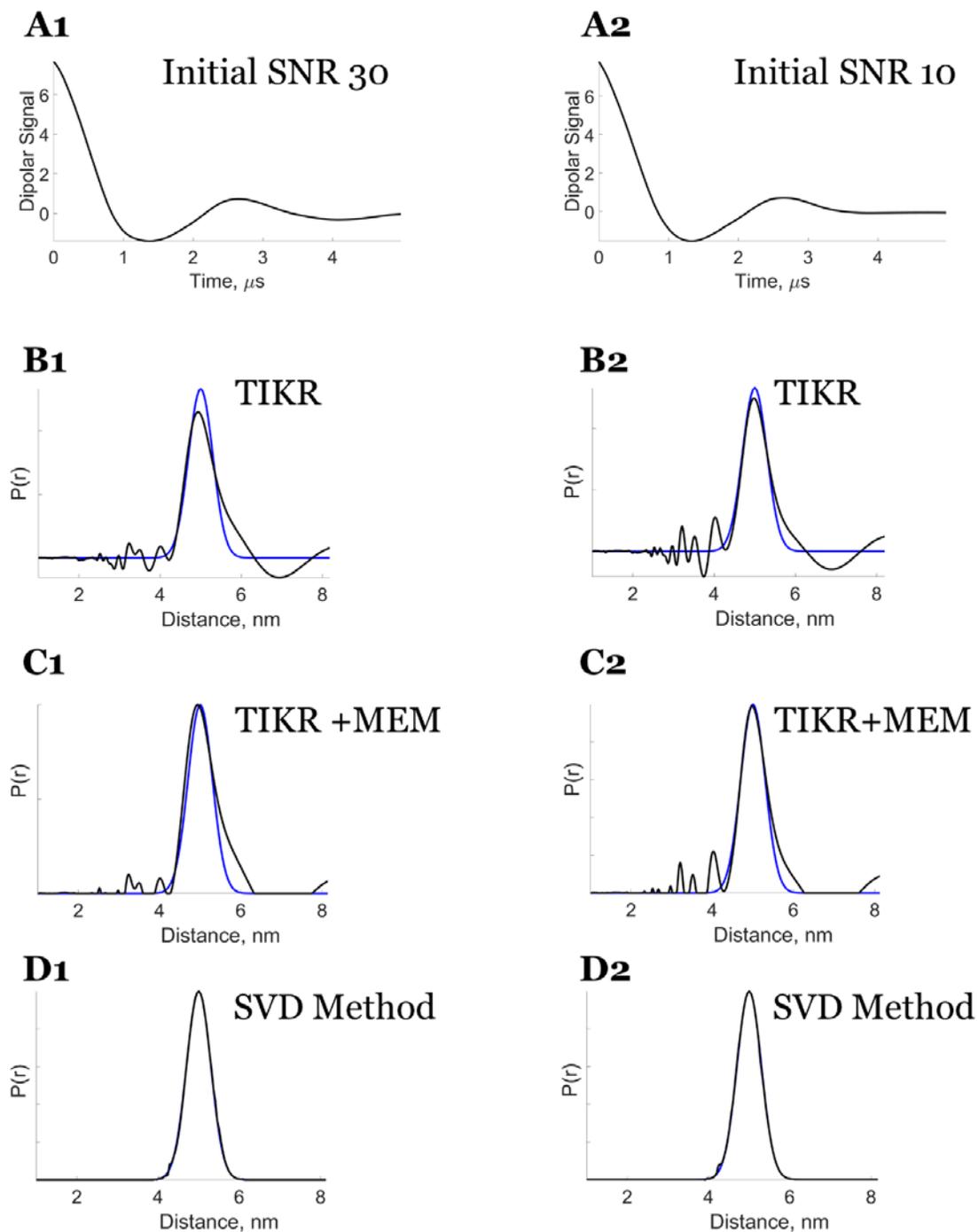
**Figure 5.S7:** Model Data with Some Noise ( $SNR \approx 850$ ) -- Unimodal Distance Distribution. **A)** Noisy model data dipolar signal; **B)** Piccard plot of the unimodal distribution from the noisy model data at different number of Singular Value Contributions (SVCs) represented by  $i$ ; and Comparison of model distribution with the distance distribution generated from: **C)**  $i = 3$ ,  $\sigma_i = 8.04$ ; **D)**  $i = 5$ ,  $\sigma_i = 3.63$ ; **E)**  $i = 8$ ,  $\sigma_i = 2.04$ ; **F)**  $i = 40$ ,  $\sigma_i = 0.33$ ; **G)**  $i = 67$ ,  $\sigma_i = 0.01$ ; **H)**  $i = 69$ ,  $\sigma_i = 3 \times 10^{-3}$ ; **I)**  $i = 71$ ,  $\sigma_i = 1 \times 10^{-3}$ ; and **J)**  $i = 82$ ,  $\sigma_i = 10^{-5}$ .

### 5.S.6 Reconstructed Bimodal Distance Distributions of Noiseless and Some Noise Model Data using new SVD method



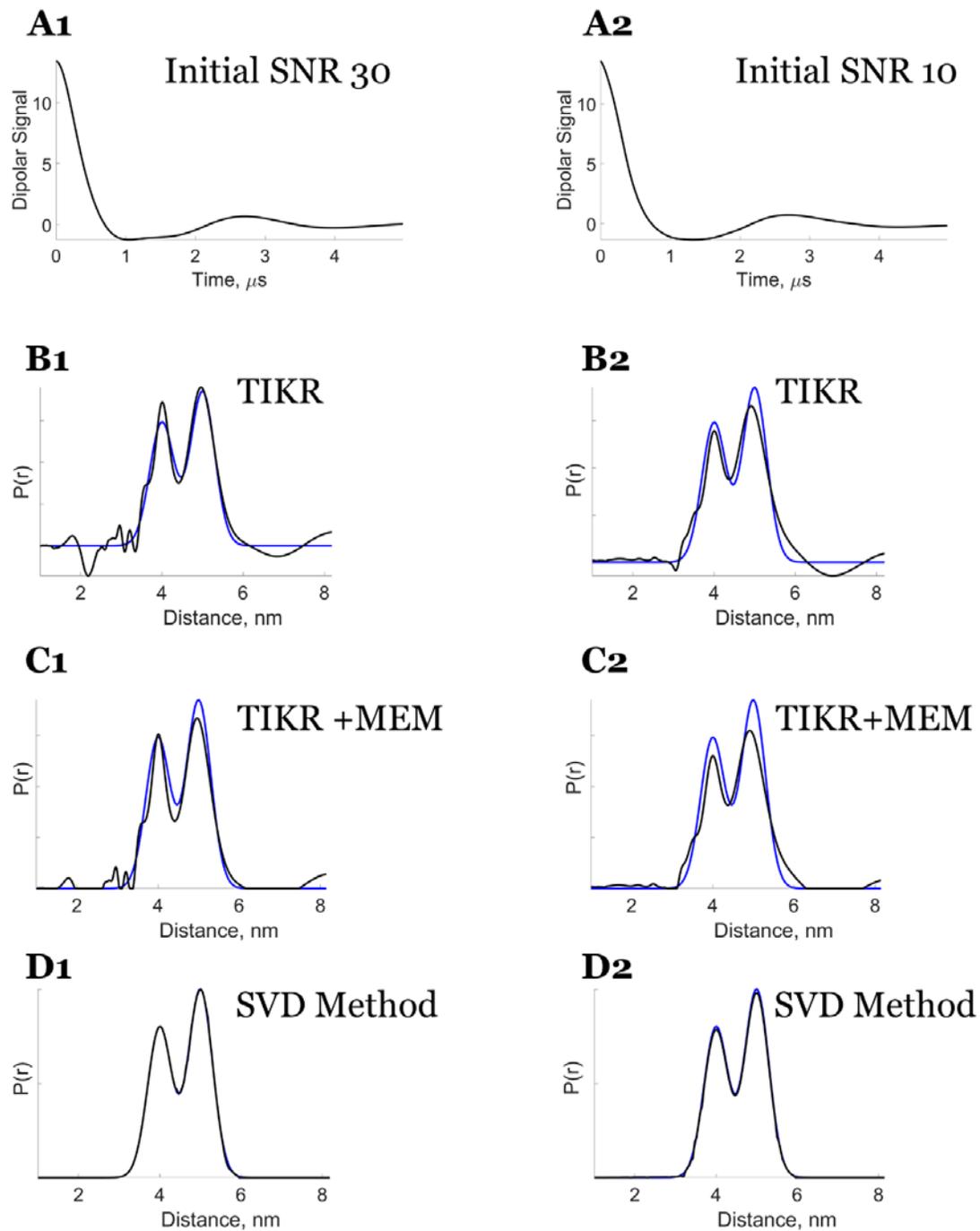
**Figure 5.S8:** Bimodal Model -- Reconstruction of distance distribution for noise-free model data and noisy model data ( $SNR \approx 850$ ) using the new SVD method. **A)** Model dipolar signal; **B)** Model dipolar signal with added noise (see added noise in Red plot); **C)** Singular value cut-off at each distance (nm) for the model dipolar signal; **D)** Singular value cut-off at each distance (nm) for the model dipolar signal with added noise; and **E)** Distance distribution reconstructed from the model dipolar signal and model dipolar signal with noise using the singular value cut-offs shown in **C)** and **D)**, respectively. Note that the added noise is so small that **A** and **B** still appear identical, but convergence to the virtually identical final results requires segmentation in the latter case.

**5.S.7 Comparison of Distance Distributions Reconstructed from TIKR, TIKR+MEM and SVD Method for Unimodal and Bimodal Dipolar Signal: Figs. 5.S9-5.S12 and Table 5.S1**



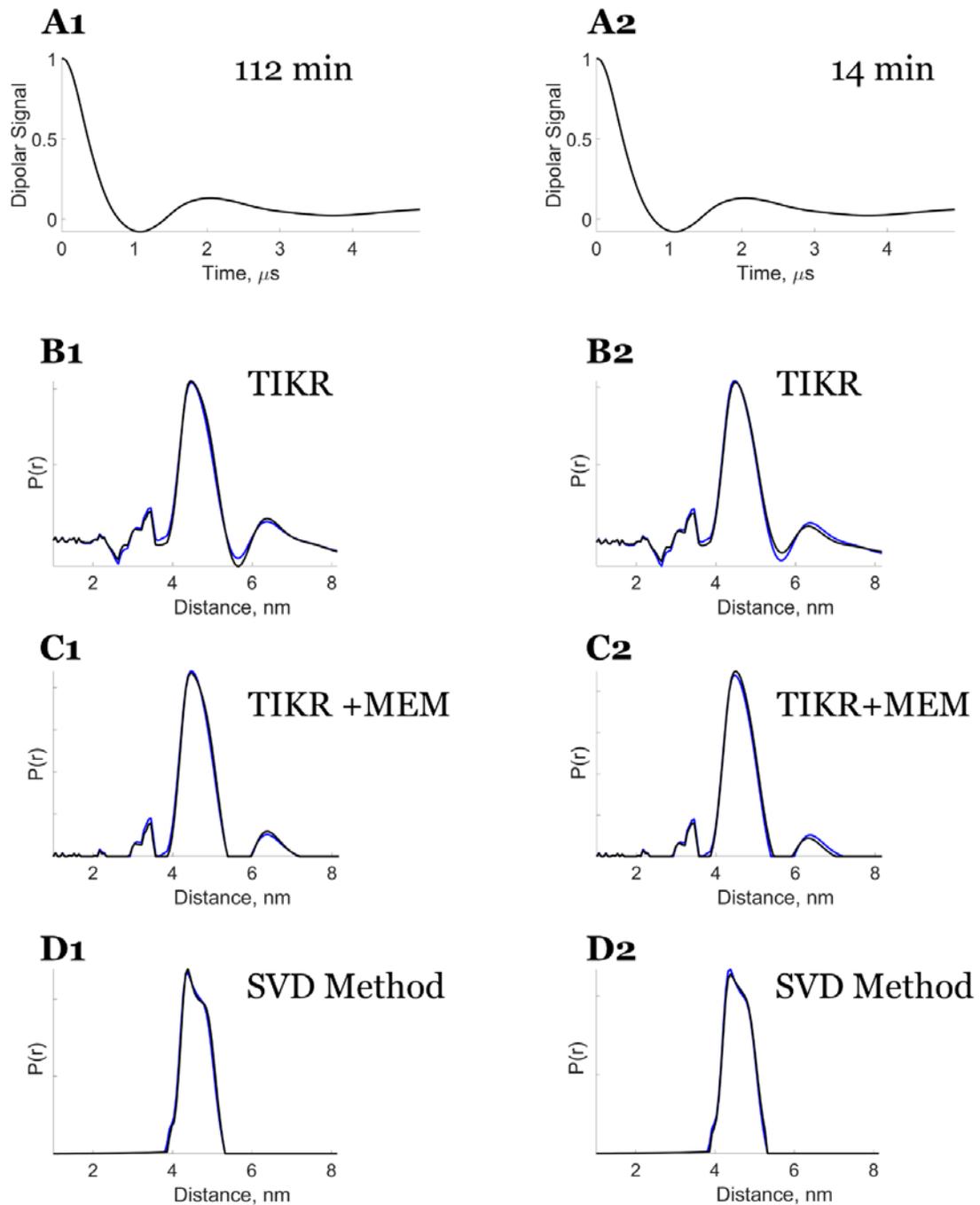
**Figure 5.S9:** Model Data -- Unimodal: Comparison of distance distributions reconstructed from Tikhonov regularization method (TIKR), Tikhonov regularization followed by maximum entropy method (TIKR+MEM), and new SVD method for model unimodal dipolar signal after first applying WavPDS denoising. **A1)** Denoised dipolar signal (Denoised  $SNR = 1880$ ) from initial noisy  $SNR = 30$ ; **B1)** Distance distribution obtained using TIKR; **C1)** Distance distribution obtained using

TIKR+MEM; **D1)** Distance distribution obtained using new SVD method; **A2)** Denoised dipolar signal (Denoised  $SNR = 850$ ) from initial noisy  $SNR = 10$ ; **B2)** Distance distribution obtained using TIKR; **C2)** Distance distribution obtained using TIKR+MEM; **D2)** Distance distribution obtained using new SVD method. Each of the distance distributions is compared with the model distance distribution (blue).



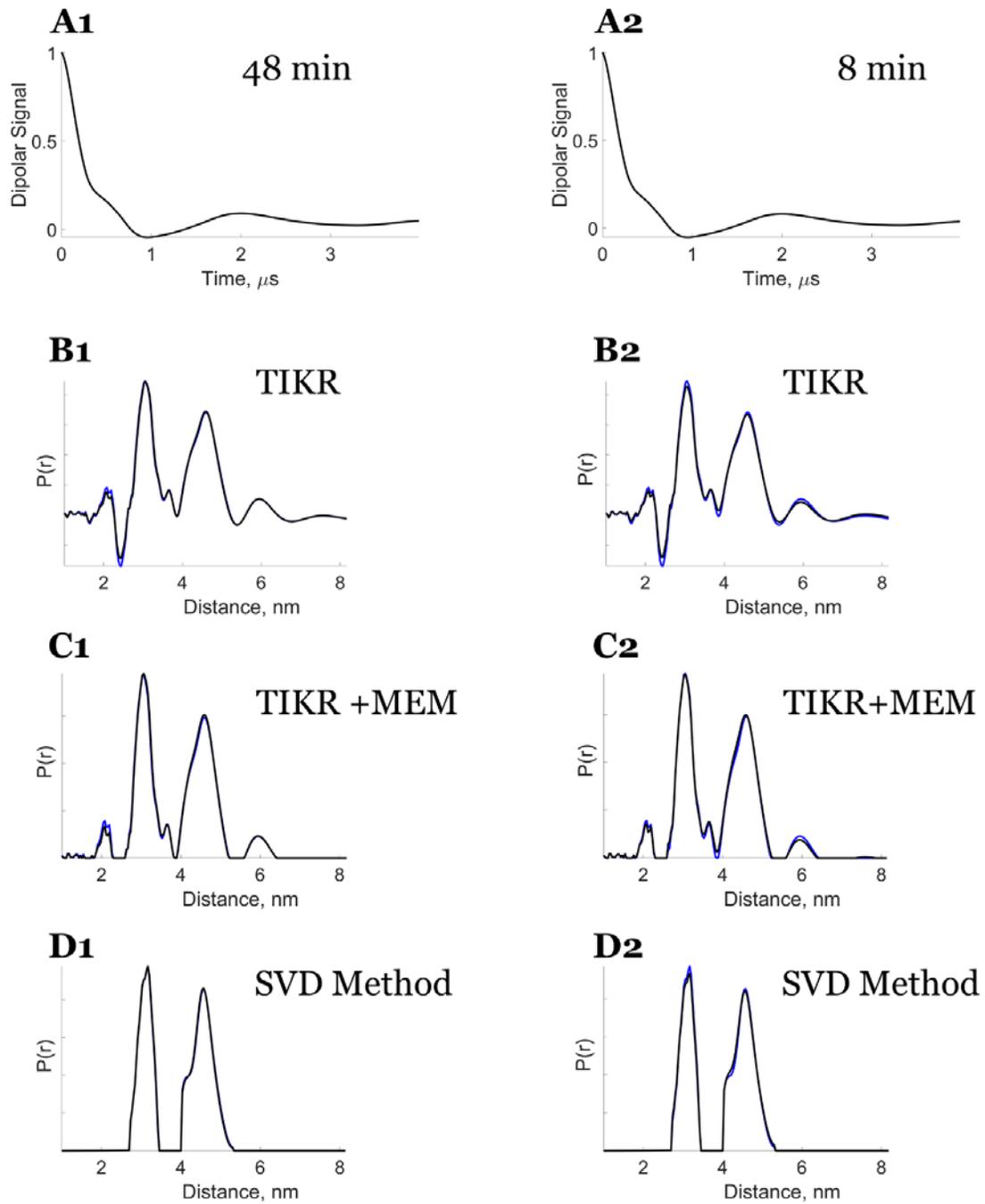
**Figure 5.S10:** Model Data -- Bimodal: Comparison of distance distributions reconstructed from Tikhonov regularization method (TIKR), Tikhonov regularization followed by maximum entropy method (TIKR+MEM), and new SVD method for model bimodal dipolar signal after first applying WavPDS denoising method. **A1)** Denoised dipolar signal (Denoised  $SNR = 3186$ ) from initial noisy  $SNR = 30$ ; **B1)**

Distance distribution obtained using TIKR; **C1)** Distance distribution obtained using TIKR+MEM; **D1)** Distance distribution obtained using new SVD method; **A2)** Denoised dipolar signal (Denoised  $SNR = 1850$ ) from initial noisy  $SNR = 10$ ; **B2)** Distance distribution obtained using TIKR; **C2)** Distance distribution obtained using TIKR+MEM; **D2)** Distance distribution obtained using new SVD method. Each of the distance distributions is compared with the model distance distribution (blue).



**Figure 5.S11:** Experimental Data -- Unimodal: Comparison of distance distributions reconstructed from Tikhonov regularization method (TIKR), Tikhonov regularization followed by maximum entropy method (TIKR+MEM), and new SVD method for experimental unimodal dipolar signal after first applying WavPDS denoising method. **A1)** Denoised dipolar signal after 112 min of signal averaging (Initial noisy  $SNR = 6.8$ , Denoised  $SNR = 909$ ); **B1)** Distance distribution obtained using TIKR; **C1)**

Distance distribution obtained using TIKR+MEM; **D1**) Distance distribution obtained using new SVD method; **A2**) Denoised dipolar signal after 14 min of signal averaging (Initial noisy  $SNR = 3.8$ , Denoised  $SNR = 488$ ); **B2**) Distance distribution obtained using TIKR; **C2**) Distance distribution obtained using TIKR+MEM; **D2**) Distance distribution obtained using new SVD method. Each of the distance distributions is compared with a distance distribution obtained from the reference signal (blue). The reference signal is denoised dipolar signal after 952 min signal averaging. The reference distance distribution was reconstructed using TIKR for **B1** and **B2**, TIKR+MEM for **C1** and **C2**, and new SVD method for **D1** and **D2**. The sample used was T4L mutant 44C/135C.<sup>14</sup>



**Figure 5.S12:** Experimental Data -- Bimodal: Comparison of distance distributions reconstructed from Tikhonov regularization method (TIKR), Tikhonov regularization followed by maximum entropy method (TIKR+MEM), and new SVD method for experimental bimodal dipolar signal after first applying WavPDS denoising method. **A1)** Denoised dipolar signal after 48 min of signal averaging (Initial noisy  $SNR = 31$ , Denoised  $SNR = 3333$ ); **B1)** Distance distribution obtained using TIKR; **C1)**

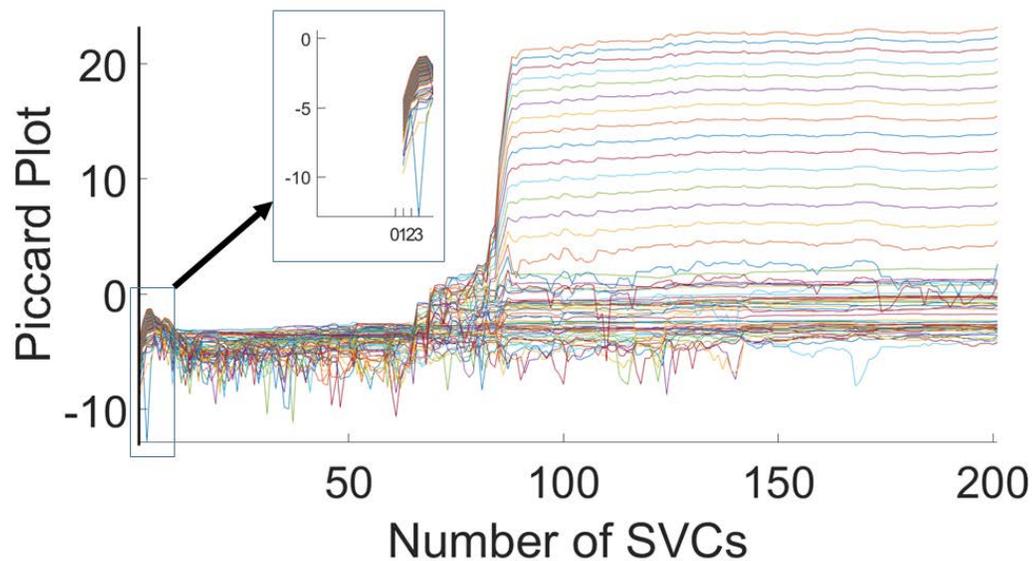
Distance distribution obtained using TIKR+MEM; **D1**) Distance distribution obtained using new SVD method; **A2**) Denoised dipolar signal after 8 min of signal averaging (Initial noisy  $SNR = 11$ , Denoised  $SNR = 1046$ ); **B2**) Distance distribution obtained using TIKR; **C2**) Distance distribution obtained using TIKR+MEM; **D2**) Distance distribution obtained using new SVD method. Each of the distance distributions is compared with a distance distribution obtained from the reference signal (blue). The reference signal is denoised dipolar signal after 952 min signal averaging. The reference distance distribution was reconstructed using TIKR for **B1** and **B2**, TIKR+MEM for **C1** and **C2**, and new SVD method for **D1** and **D2**. The sample used consisted of T4L mutants 8C/44C and 44C/135C.

**Table 5.S1:** Signal-to-Noise Ratios (SNRs) of the unimodal and bimodal dipolar signals for model and experimental cases before and after denoising.

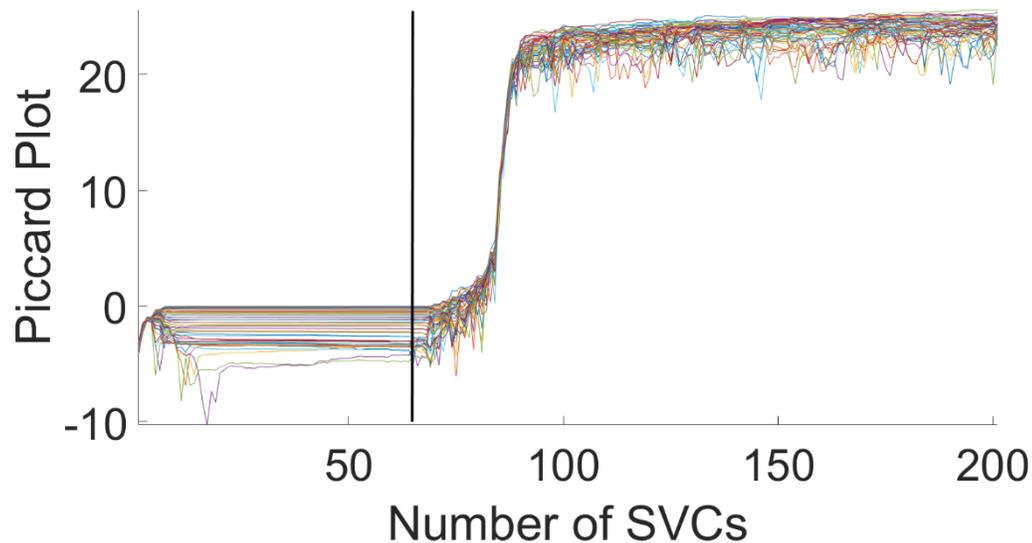
<b>Dipolar Signal</b>	<b>Initial SNR</b>	<b>Denoised SNR</b>
Model – Unimodal (cf. Fig. 5.S9)	30	1880
	10	850
Model – Bimodal (cf. Fig. 5.S10)	30	3186
	10	1850
Experimental – Unimodal (cf. Fig. 5.S11)	6.8	909
	3.8	488
Experimental – Bimodal (cf. Fig. 5.S12)	31	3333
	11	1046

### 5.S.8 Piccard Plots for Individual $r$ -values at Each Segmented Region:

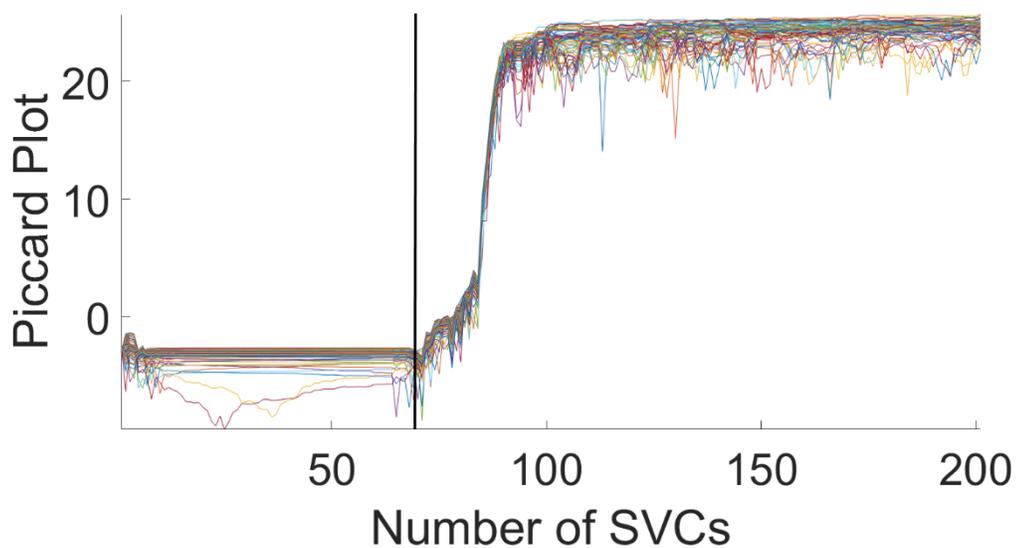
Figs. 5.S13-5.S20



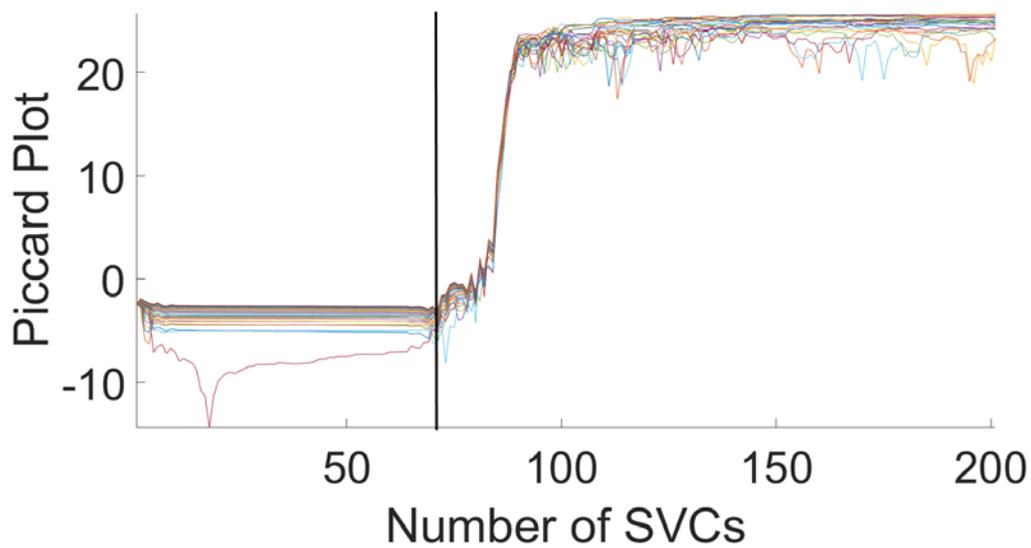
**Figure 5.S13:** Model Data – Unimodal: Singular value cut-off on the Piccard plot for segment 1 (1 to 3.88 nm) of  $r$  region.



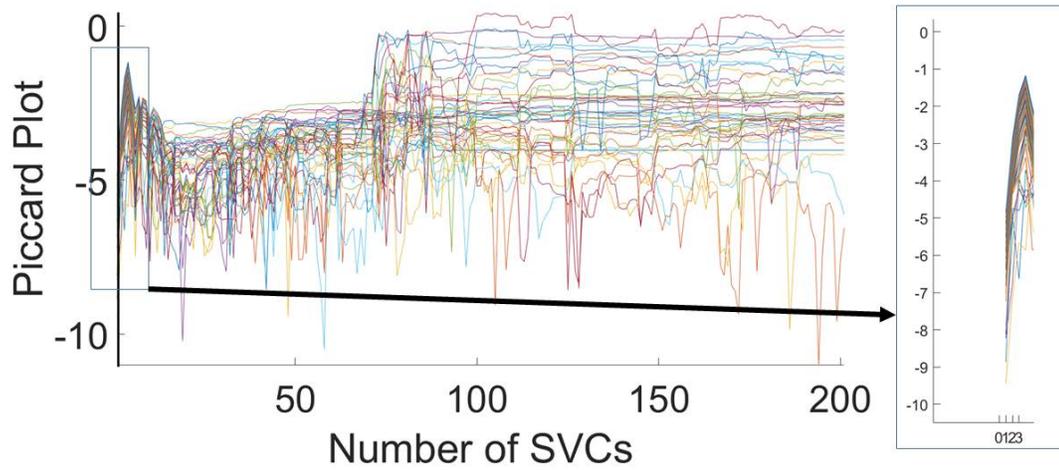
**Figure 5.S14:** Model Data – Unimodal: Singular value cut-off on the Piccard plot for segment 2 (3.92 to 6.08 nm) of  $r$  region.



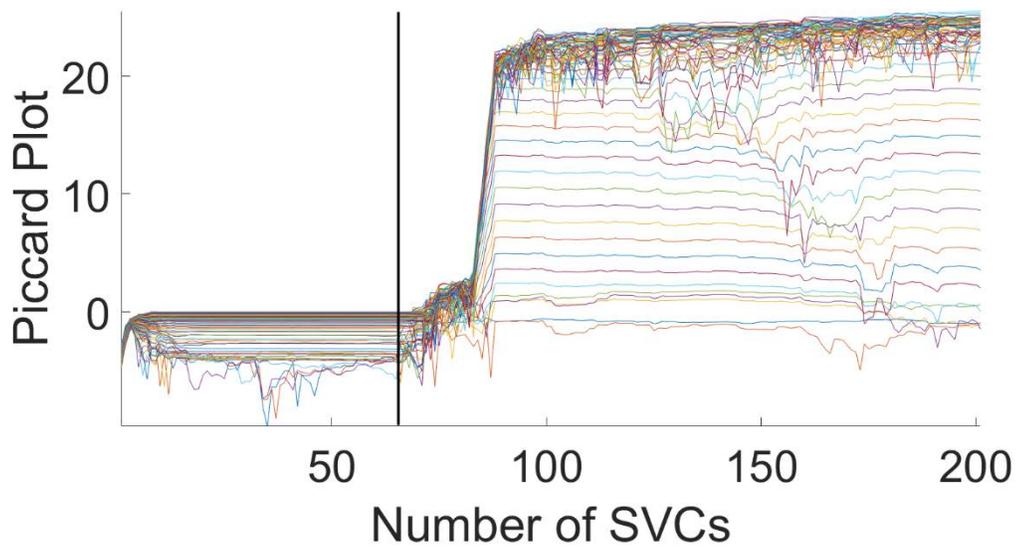
**Figure 5.S15:** Model Data – Unimodal: Singular value cut-off on the Piccard plot for segment 2 (6.13 to 8.96 nm) of  $r$  region.



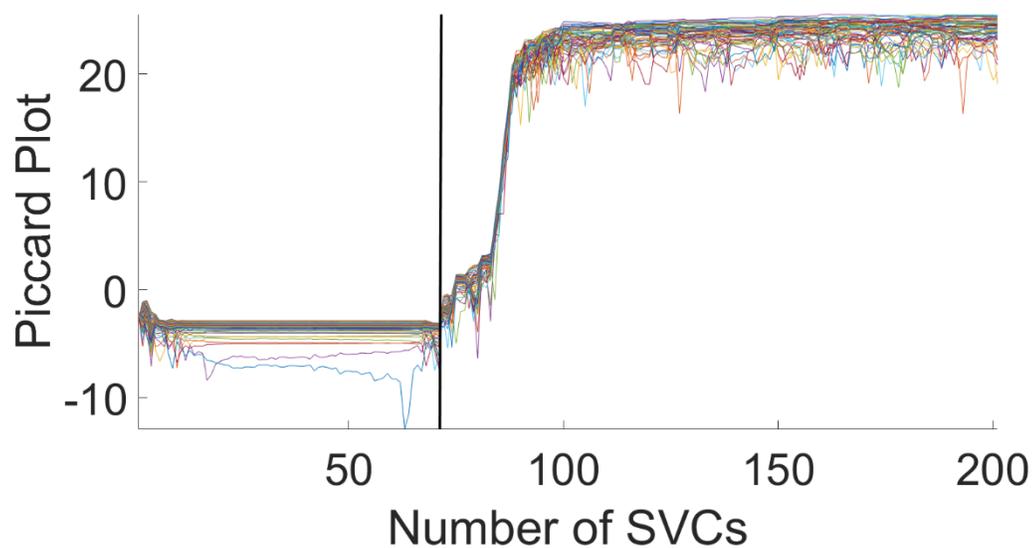
**Figure 5.S16:** Model Data – Unimodal: Singular value cut-off on the Piccard plot for segment 2 (9.01 to 10 nm) of  $r$  region.



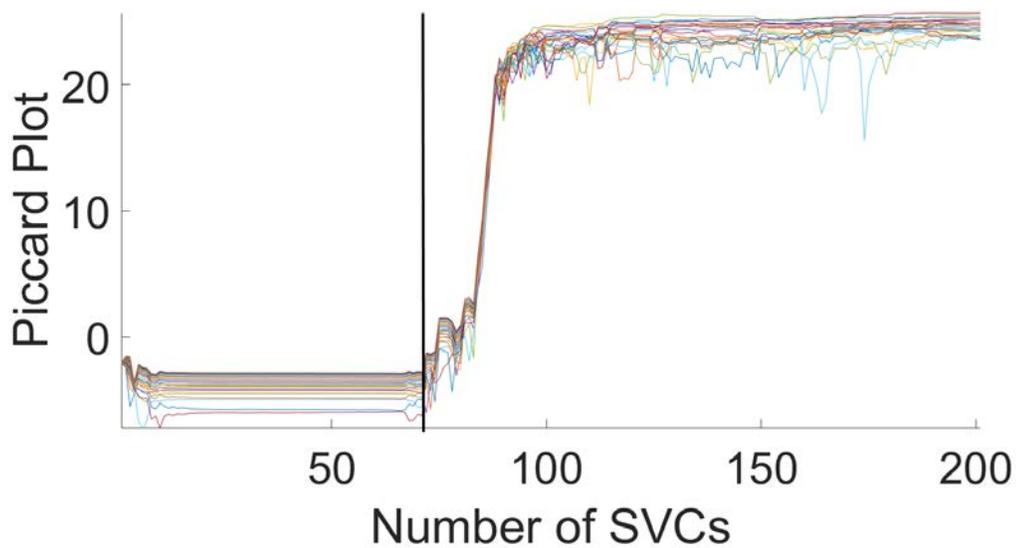
**Figure 5.S17:** Model Data – Bimodal: Singular value cut-off on the Piccard plot for segment 1 (1 to 2.89 nm) of  $r$  region.



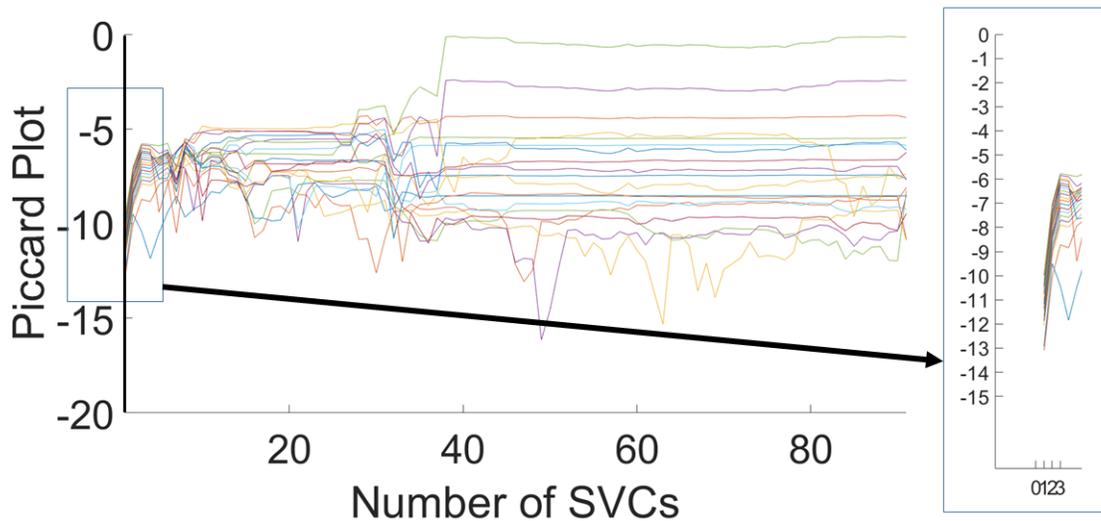
**Figure 5.S18:** Model Data – Bimodal: Singular value cut-off on the Piccard plot for segment 2 (2.93 to 6.08 nm) of  $r$  region.



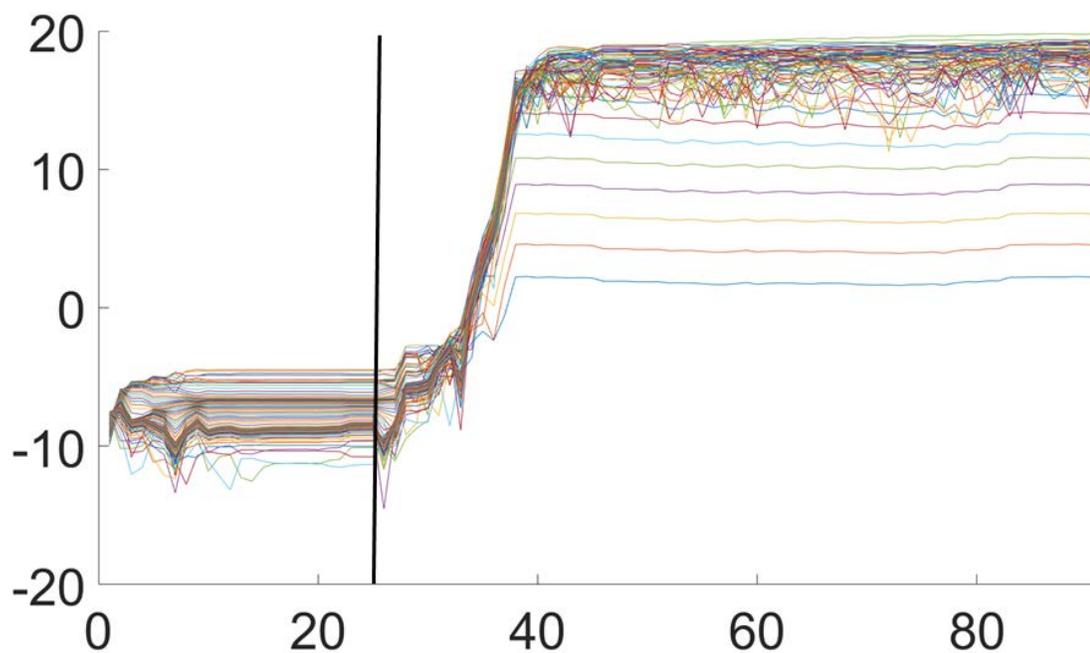
**Figure 5.S19:** Model Data – Bimodal: Singular value cut-off on the Piccard plot for segment 3 (6.13 to 8.96 nm) of  $r$  region.



**Figure 5.S20:** Model Data – Bimodal: Singular value cut-off on the Piccard plot for segment 4 (9.01 to 10 nm) of  $r$  region.



**Figure 5.S21:** Experimental Data (cf. Fig. 4): Singular value cut-off on the Piccard plot for segment 1 (1 to 2.7 nm) of  $r$  region.



**Figure 5.S22:** Experimental Data (cf. Fig. 4): Singular value cut-off on the Piccard plot for segment 2 (2.8 to 10 nm) of  $r$  region.

## References

- (1) Allen, G. I.; Peterson, C.; Vannucci, M.; Maletić-Savatić, M. Regularized Partial Least Squares with an Application to NMR Spectroscopy. *Stat. Anal. Data Min.* **2013**, *6* (4), 302–314.
- (2) Borbat, P. P.; Freed, J. H. Pulse Dipolar Electron Spin Resonance: Distance Measurements. In *Structure and Bonding*; Springer Berlin Heidelberg, 2013; pp 1–82.
- (3) Kuramshina, G. M.; Weinhold, F.; Kochikov, I. V.; Yagola, A. G.; Pentin, Y. A. Joint Treatment of Ab Initio and Experimental Data in Molecular Force Field Calculations with Tikhonov's Method of Regularization. *J. Chem. Phys.* **1994**, *100* (2), 1414–1424.
- (4) Fidêncio, P. H.; Poppi, R. J.; de Andrade, J. C. Determination of Organic Matter in Soils Using Radial Basis Function Networks and near Infrared Spectroscopy. *Anal. Chim. Acta* **2002**, *453* (1), 125–134.
- (5) Tyler, B. J.; Castner, D. G.; Ratner, B. D. Regularization: A Stable and Accurate Method for Generating Depth Profiles from Angle-Dependent XPS Data. *Surf. Interface Anal.* **1989**, *14* (8), 443–450.
- (6) Hasekamp, O. P.; Landgraf, J. Ozone Profile Retrieval from Backscattered Ultraviolet Radiances: The Inverse Problem Solved by Regularization. *J. Geophys. Res. Atmos.* **2001**, *106* (D8), 8077–8088.
- (7) Chiang, Y.-W.; Borbat, P. P.; Freed, J. H. The Determination of Pair Distance Distributions by Pulsed ESR Using Tikhonov Regularization. *J. Magn. Reson.* **2005**, *172* (2), 279–295.
- (8) Chiang, Y.-W.; Borbat, P. P.; Freed, J. H. Maximum Entropy: A Complement to Tikhonov Regularization for Determination of Pair Distance Distributions by Pulsed ESR. *J. Magn. Reson.* **2005**, *177* (2), 184–196.
- (9) Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J.; Timmel, C. R.; Hilger, D.; Jung, H. DeerAnalysis2006—a Comprehensive Software Package for Analyzing Pulsed ELDOR Data. *Appl. Magn. Reson.* **2006**, *30* (3–4), 473–498.

- (10) Jeschke, G.; Panek, G.; Godt, A.; Bender, A.; Paulsen, H. Data Analysis Procedures for Pulse ELDOR Measurements of Broad Distance Distributions. *Appl. Magn. Reson.* **2004**, *26* (1–2), 223–244.
- (11) Edwards, T. H.; Stoll, S. A Bayesian Approach to Quantifying Uncertainty from Experimental Noise in DEER Spectroscopy. *J. Magn. Reson.* **2016**, *270*, 87–97.
- (12) Horn, R. A.; Johnson, C. A. *Matrix Analysis*, First Edit.; Cambridge University Press: New York, 1985.
- (13) Srivastava, M.; Anderson, C. L.; Freed, J. H. A New Wavelet Denoising Method for Selecting Decomposition Levels and Noise Thresholds. *IEEE Access* **2016**, *4*, 3862–3877.
- (14) Srivastava, M.; Georgieva, E. R.; Freed, J. H. A New Wavelet Denoising Method for Experimental Time-Domain Signals: Pulsed Dipolar Electron Spin Resonance. *J. Phys. Chem. A* **2017**, *121* (12), 2452–2465.
- (15) Borbat, P. P.; Freed, J. H. Measuring Distances by Pulsed Dipolar ESR Spectroscopy: Spin-Labeled Histidine Kinases. *Methods Enzymol.* **2007**, *423*, 52–116.
- (16) Georgieva, E. R.; Borbat, P. P.; Ginter, C.; Freed, J. H.; Boudker, O. Conformational Ensemble of the Sodium-Coupled Aspartate Transporter. *Nat. Struct. Mol. Biol.* **2013**, *20* (2), 215–221.
- (17) Jeschke, G. DEER Distance Measurements on Proteins. *Annu. Rev. Phys. Chem.* **2012**, *63* (1), 419–446.
- (18) Klare, J. P.; Steinhoff, H.-J. Structural Information from Spin-Labelled Membrane-Bound Proteins; Springer Berlin Heidelberg, 2013; pp 205–248.
- (19) Mchaourab, H. S.; Steed, P. R.; Kazmier, K. Toward the Fourth Dimension of Membrane Protein Structure: Insight into Dynamics from Spin-Labeling EPR Spectroscopy. *Structure* **2011**, *19* (11), 1549–1561.
- (20) Edwards, D. T.; Huber, T.; Hussain, S.; Stone, K. M.; Kinnebrew, M.; Kaminker, I.; Matalon, E.; Sherwin, M. S.; Goldfarb, D.; Han, S. Determining

the Oligomeric Structure of Proteorhodopsin by Gd<sup>3+</sup>-Based Pulsed Dipolar Spectroscopy of Multiple Distances. *Structure* **2014**, 22 (11), 1677–1686.

- (21) DeBerg, H. A.; Bankston, J. R.; Rosenbaum, J. C.; Brzovic, P. S.; Zagotta, W. N.; Stoll, S. Structural Mechanism for the Regulation of HCN Ion Channels by the Accessory Protein TRIP8b. *Structure* **2015**, 23 (4), 734–744.
- (22) Schiemann, O.; Prisner, T. F. Long-Range Distance Determinations in Biomacromolecules by EPR Spectroscopy. *Q. Rev. Biophys.* **2007**, 40 (01), 1.
- (23) Borbat, P. P.; Mchaourab, H. S.; Freed, J. H. Protein Structure Determination Using Long-Distance Constraints from Double-Quantum Coherence ESR: Study of T4 Lysozyme. *J. Am. Chem. Soc.* **2002**, 124 (19), 5304–5314.
- (24) Hansen, P. C.; O’Leary, D. P. The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems. *SIAM J. Sci. Comput.* **1993**, 14 (6), 1487–1503.
- (25) Raitsimring, A.; Salikhov, K. Electron Spin Echo Method as Used to Analyze the Spatial Distribution of Paramagnetic Centers. *Bull Magn Reson* **1985**, 7 (4), 184–217.
- (26) Milov, A. D.; Maryasov, A. G.; Tsvetkov, Y. D. Pulsed Electron Double Resonance (PELDOR) and Its Applications in Free-Radicals Research. *Appl. Magn. Reson.* **1998**, 15 (1), 107–143.
- (27) Milov, A. D.; Tsvetkov, Y. D.; Formaggio, F.; Oancea, S.; Toniolo, C.; Raap, J. Aggregation of Spin Labeled Trichogin GA IV Dimers: Distance Distribution between Spin Labels in Frozen Solutions by PELDOR Data. *J. Phys. Chem. B* **2003**, 107 (49), 13719–13727.
- (28) Stein, R. A.; Beth, A. H.; Hustedt, E. J. A Straightforward Approach to the Analysis of Double Electron–Electron Resonance Data. In *Methods in enzymology*; 2015; Vol. 563, pp 531–567.
- (29) Dalmas, O.; Hyde, H. C.; Hulse, R. E.; Perozo, E. Symmetry-Constrained Analysis of Pulsed Double Electron–Electron Resonance (DEER) Spectroscopy Reveals the Dynamic Nature of the KcsA Activation Gate. *J. Am. Chem. Soc.* **2012**, 134 (39), 16360–16369.
- (30) Pannier, M.; Schädler, V.; Schöps, M.; Wiesner, U.; Jeschke, G.; Spiess, H. W.

Determination of Ion Cluster Sizes and Cluster-to-Cluster Distances in Ionomers by Four-Pulse Double Electron Electron Resonance Spectroscopy. *Macromolecules* **2000**, *33* (21), 7812–7818.

- (31) Sen, K. I.; Logan, T. M.; Fajer, P. G. Protein Dynamics and Monomer–Monomer Interactions in AntR Activation by Electron Paramagnetic Resonance and Double Electron–Electron Resonance †. *Biochemistry* **2007**, *46* (41), 11639–11649.
- (32) Masureel, M.; Martens, C.; Stein, R. A.; Mishra, S.; Ruyschaert, J.-M.; Mchaourab, H. S.; Govaerts, C. Protonation Drives the Conformational Switch in the Multidrug Transporter LmrP. *Nat. Chem. Biol.* **2013**, *10* (2), 149–155.
- (33) Fajer, P. G.; Brown, L.; Song, L. Practical Pulsed Dipolar ESR (DEER). In *ESR Spectroscopy in Membrane Biophysics*; Springer US: Boston, MA, 2007; pp 95–128.
- (34) Brandon, S.; Beth, A. H.; Hustedt, E. J. The Global Analysis of DEER Data. *J. Magn. Reson.* **2012**, *218*, 93–104.
- (35) Matveeva, A. G.; Yushkova, Y. V.; Morozov, S. V.; Grygor'ev, I. A.; Dzuba, S. A. Multi-Gaussian Monte Carlo Analysis of PELDOR Data in the Frequency Domain. *Zeitschrift für Phys. Chemie* **2017**, *231* (3).
- (36) Srivastava, M.; Georgieva, E. R.; Freed, J. H. A New Wavelet Denoising Method for Experimental Time-Domain Signals: Pulsed Dipolar Electron Spin Resonance. *J. Phys. Chem. A* **2017**, *121* (12), 2452–2465.
- (37) Srivastava, M.; Anderson, C. L.; Freed, J. H. A New Wavelet Denoising Method for Selecting Decomposition Levels and Noise Thresholds. *IEEE Access* **2016**, *4*, 3862–3877.
- (38) Golub, G. . H.; Van Loan, C. . F. *Matrix Computations*, 4th ed.; Johns Hopkins University Press: Baltimore, MD, 2013.
- (39) Groetsch, C. W. *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*; Pitman Advanced Publication Program: Boston, 1984.
- (40) Hansen, P. C. *Analysis of Discrete Ill-Posed Problems by Means of the L-*

Curve. *SIAM Rev.* **1992**, 34 (4), 561–580.

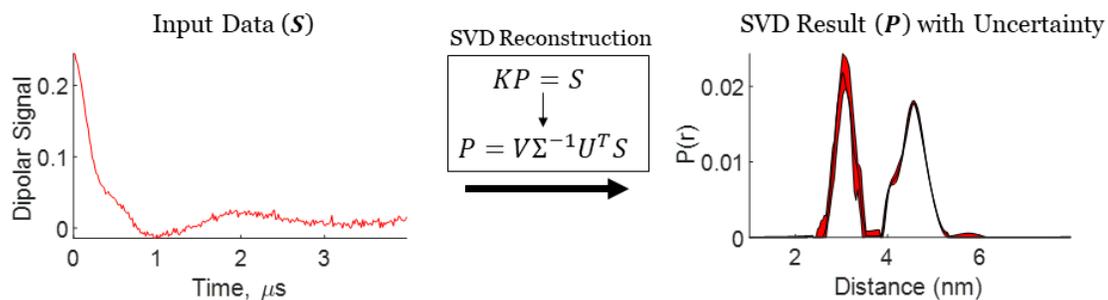
# **CHAPTER 6**

## **SINGULAR VALUE DECOMPOSITION METHOD TO DETERMINE DISTANCE DISTRIBUTIONS IN PULSE DIPOLAR ELECTRON SPIN RESONANCE: II. ESTIMATING UNCERTAINTY**

### **Abstract**

This chapter continues Srivastava and Freed (2017) on using singular value decomposition (SVD) to obtain physical results from experimental signals, without any need for regularization or other methods limited in applicability. We show here how to calculate uncertainty in SVD-generated solutions. With sufficiently high Signal-to-Noise Ratio, the uncertainty in the solution is obtained by finding the minimum and maximum values for the region where the solution, obtained from a different range of singular value contributions, remains converged. The width of the converged region varies based on the noise level and data point location, resulting in well-resolved uncertainties. The uncertainty levels typically found are very small, emphasizing the reliability of the method. Pulsed dipolar electron spin resonance spectroscopy experiments are used as an example, but the SVD approach is general and thus applicable to any similar method.

## Graphical Abstract



\*This Chapter will be submitted for publication.

**Authors:** Madhur Srivastava, and Jack H. Freed

**Author Contributions:** Madhur Srivastava conducted the research and wrote the paper, and Jack Freed supervised the research and edited the paper.

## 6.1 Introduction

In many physical systems, methods based on singular value decomposition (SVD) are used to process data to elicit desired results.<sup>1-8</sup> However, uncertainty determination in the resultant outcome has remained a challenge and is needed to reliably interpret the results. There have been some efforts to determine uncertainty, but they focused on from the overall outcome<sup>9,10</sup> effects of noise present in the experimental signal and on other factors such as background signal<sup>10</sup>. Improvements in instrument technology<sup>2,11-14</sup> and developments in data processing methods, such as wavelet denoising<sup>15,16</sup>, have enabled obtaining experimental signals with very high Signal-to-Noise Ratio (SNR). This shifts the focus from uncertainty due to the noise to the data processing using SVD. The uncertainty generated due to the SVD process has received little attention but is considered here.

We describe a method to quantify the uncertainty resulting from the SVD-based reconstruction of physical experiments. Our previous work has shown that the SVD solution at each output point can converge for a different number of singular value contributions (SVCs) before becoming unstable.<sup>7</sup> This uncertainty in convergence at each output point is not uniform and varies between such points. Thus, we determine the uncertainty at each location of the output independently, based on the SVCs.

We demonstrate the analysis using pulsed dipolar electron spin resonance spectroscopy (PDS) as an example. PDS plays a key role in determining structure and dynamics of biological systems.<sup>2,17-24</sup>

## 6.2 Method

### 6.2.1 Pulsed Dipolar Spectroscopy

In PDS, paramagnetic tags<sup>25,26</sup> called spin labels (NO, Cu<sup>2+</sup>, among others) are attached at specific locations and then a dipolar signal is acquired from the interaction between a pair of spin labels, from which the distance distribution between them,  $P(r)$  may be obtained. The distances between them can then be used to determine the number of conformations in a biomolecule and their relative populations. The distances may range between 1 to 10 nm.

### 6.2.2 Distance Distribution Reconstruction in PDS using SVD

The distance distributions  $P(r)$  are constructed using a mathematical inversion process involving the dipolar signal  $S$  and a kernel matrix  $K$ . The mathematical representation in matrix notation is as follows,

$$KP = S \tag{6.1}$$

The dipolar signal  $S(t)$  is acquired from the experiment, while the kernel matrix  $K$  contains the dipolar interaction between spin pairs and is generated from theory as a function of time and distance. The matrix dimensions of  $K$ ,  $P$ , and  $S$  are  $M \times N$ ,  $N \times 1$ , and  $M \times 1$ , respectively, where  $N \leq M$ . One wishes to obtain the distance distribution  $P$  given  $K$  and  $S$ . Since  $K$  is a singular matrix, mathematical inversion ( $P = K^{-1}S$ ) cannot be obtained simply. Due to the ill-posed nature of the problem, the

SVD method is used to carry out the inversion. The  $K$  matrix can be decomposed in the SVD form as,

$$K = U\Sigma V^T \quad (6.2)$$

where  $U$  and  $V$  are orthogonal matrices with dimensions  $M \times M$  and  $N \times N$ , respectively, and  $\Sigma$  is an  $M \times N$  diagonal matrix consisting of non-negative singular values in decreasing order. The singular values in  $\Sigma$  are denoted as  $\sigma_i$  (where  $i = 1, 2, \dots, N$ ). Using equation 6.2, equation 1 can be rewritten as,

$$U\Sigma V^T P = S \quad (6.3)$$

The dipolar signal  $P$  can be obtained from equation 3 using pseudo-inversion as follows,

$$P = V\Sigma^{-1}U^T S \quad (6.4)$$

where  $\Sigma^{-1}$  now contains the reciprocal of just the non-zero singular values in the diagonal matrix, i.e. the  $\left(\frac{1}{\sigma_i}\right)$ . The SVD-based solution<sup>7</sup> for  $P$  is given by,

$$P_j = \sum_{i=1}^{k_j} \sum_{l=1}^M V_{ji} (\Sigma^{-1})_{ii} U^T_{il} S_l \quad (6.5)$$

where  $j$ ,  $i$  and  $l$  are row and column indices, and  $k$  is the number of non-zero singular values. For physical interpretation,  $j$  is the index that represents the discretized distance,  $l$  represents the discretized time, and  $i$  is the index representing the particular singular value.

The solution of  $P_j$  should be obtained for each distance value,  $j$  as given in equation 5. This is because, for signals with noise, each  $P_j$  converges at a different singular value cut-off  $k_j$  before becoming unstable. One can select a single singular value cut-off  $k_j$ , but it would yield a good resolution over some distance ranges but unstable distributions over other ranges.

### 6.2.3 Uncertainty Determination

The uncertainty of the distribution for each distance or distance range is determined using the following three steps:

#### 6.2.3.1 Determination of Singular Value Cut-off

The singular value cut-off is determined for each distance or distance range after which the distribution value becomes unstable. This is accomplished using the Piccard condition, which is written as,

$$[Piccard]_j = \log_{10} \left( \sum_{i=1}^{c_j} \left( \frac{U^T_i S V_j}{\sigma_i} \right)^2 \right) < \infty$$

(6.6)

where  $l \leq k < M$ . The Piccard condition informs about the convergence of the solution. If the denominator  $(\sigma_i)^2$  for the  $i^{th}$  singular value decays sufficiently faster than the numerator  $(U^T_i S V_j)^2$ , the condition diverges resulting in an unstable solution.  $c_j$  is taken as the last singular value,  $\sigma_i$  for which the solution remains converged. The  $c_j$  value can be obtained with the help of a Piccard plot that reveals the point at which the Piccard condition is violated. Figs. 6.1B and 6.1C show Piccard plots at 2 nm and 4.3 nm for a unimodal distribution, Fig. 6.1A.

### **6.2.3.2 Identification of Minimum and Maximum SVCs**

The SVCs at which the distribution has converged and remains converged are determined. It is accomplished by finding the minimum and maximum SVC values of the converged region. The minimum and maximum SVCs can be found with the help of a Piccard plot or by plotting the  $P(r)$  values w.r.t. SVCs. The SVCs in the plot are restricted from 1 to  $c_j$  because after  $c_j$  the SVC values diverge. The location of first and last values of  $j$  can be denoted as  $Min_j$  and  $Max_j$ , respectively. Figs. 6.1D and 6.1E show the values of the distribution at 2 nm and 4.3 nm for a unimodal distribution, revealing the converged range.

### 6.2.3.3 Calculation of Minimum and Maximum Value of the Distance in the Min-Max SVC Range

The minimum  $P_j^{min}$  and maximum  $P_j^{max}$  value of the distribution in the converged region is between  $Min_j$  and  $Max_j$  and given as follows,

$$P_j^{min} = \min\{P_{ji}: Min_j \leq i \leq Max_j\} \quad (6.7)$$

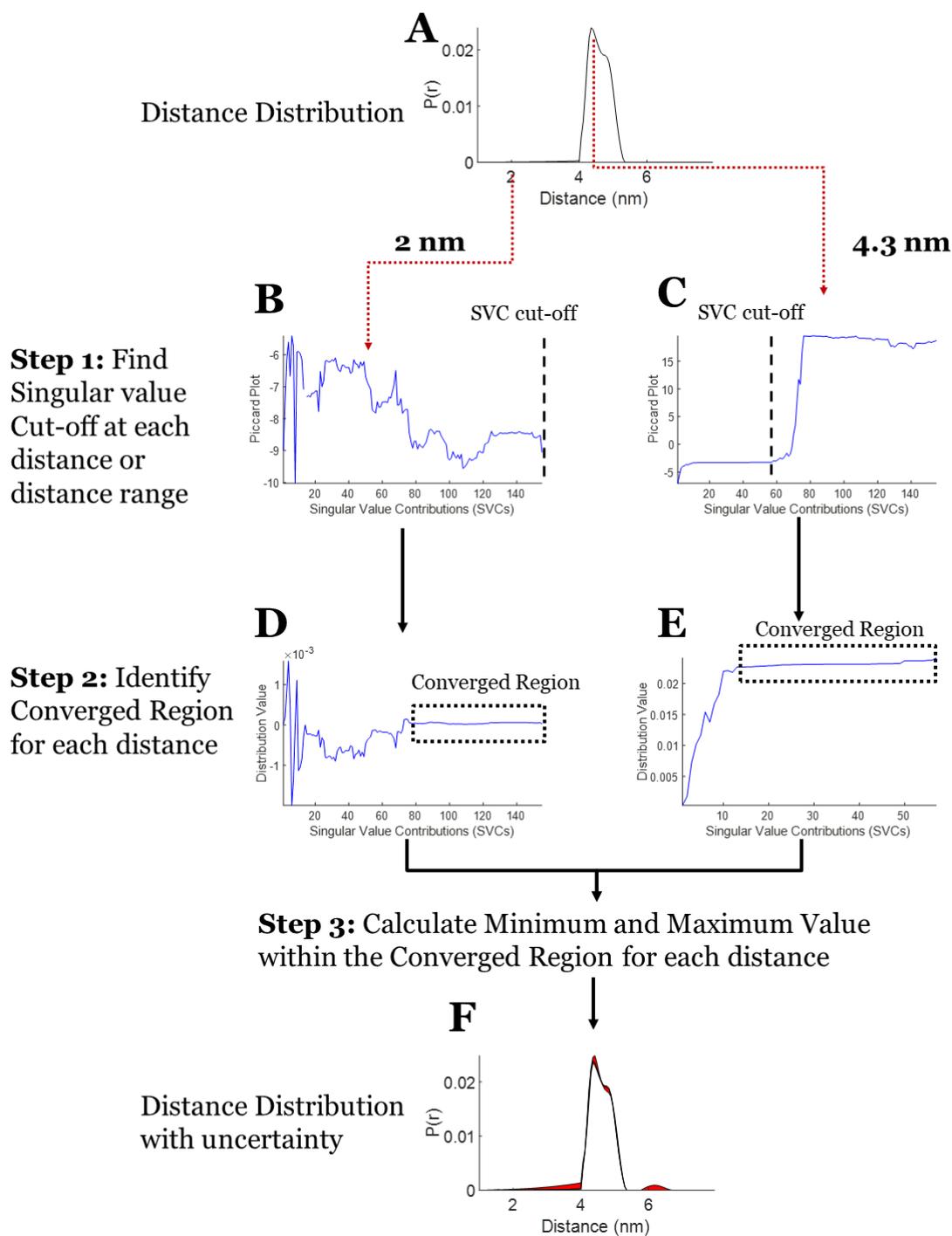
and,

$$P_j^{max} = \max\{P_{ji}: Min_j \leq i \leq Max_j\} \quad (6.8)$$

The  $P_j^{min}$  and  $P_j^{max}$  form the uncertainty range.

The optimal distance value  $P_j^{Opt}$  is given as,

$$P_j^{Min} \leq P_j^{Opt} \leq P_j^{Max} \quad (6.9)$$



**Figure 6.1:** The step-wise process of obtaining uncertainty in the distance distribution from the dipolar signal. **A)** Distance distribution reconstructed using the SVD method<sup>7</sup>; **B)** Piccard plot for the 2 nm distance, revealing that the solution never diverges. Hence, the last data point is selected as SVC cut-off; **C)** Piccard plot for the 4.3 nm distance, revealing the singular value cut-off before which solution at the

distance diverges; **D)**  $P(r)$  values obtained at the 2 *nm* distance by different SVCs until the SVC cut-off; if it never diverges than last SVC is selected as cut-off **E)**  $P(r)$  values obtained at the 4.3 *nm* distance by different SVCs until the SVC cut-off; and **F)** Distance distribution with uncertainty in distribution shown in red.

## 6.3 Results and Discussion

### 6.3.1 Experimental Data

Three sets of experimental data are used to illustrate the uncertainty analysis. Case 1 (cf. Fig. 6.2) consists of a unimodal distribution from T4 lysozyme (T4L) where the 63  $\mu\text{M}$  protein sample is of spin-labeled T4L mutant 44C/135C. The signal was averaged for 952 min to obtain a *SNR* of 37. In case 2 (cf. Fig 6.3), a bimodal distribution was used that is also generated from the T4L, where the sample is a mixture of mutants spin-labeled at 8C/44C and 44C/135C with concentrations of 44 $\mu\text{M}$  and 47 $\mu\text{M}$ , respectively. The dipolar signal was acquired after signal averaging for 360 min to obtain a *SNR* of 80. The data from the samples are the same as used in both the WavPDS method paper<sup>16</sup> and the SVD paper<sup>7</sup>. Case 3 (cf. Fig. 6.4) is also from a sample containing a biomodal distribution but at a sub-micromolar concentration. The sample is a spin-labeled Immunoglobulin E (IgE) cross-linked with trivalent DNA-DNP ligand in Phosphate Buffered Saline (PBS) buffer<sup>7,27</sup>. The dipolar signal was obtained after 18-hour of signal averaging resulting in a *SNR* of 3.8. The WavPDS method is then applied to remove the noise from the dipolar signals in all three cases (cf. Figs. 6.2B, 6.3B, 6.4C).

In cases 1 and 2, WavPDS was applied after baseline subtraction, whereas in case 3, the denoising is applied before baseline subtraction in the log-domain (cf. Fig. 6.4A). The denoising greatly helped to obtain an accurate baseline in case 3, as it was not

possible from the noisy data with  $SNR$  of 3.8. The baseline to subtract obtained from the denoised data was used for both the noisy and denoised data in this illustration.

### 6.3.2 Distance Reconstruction and Uncertainty Analysis

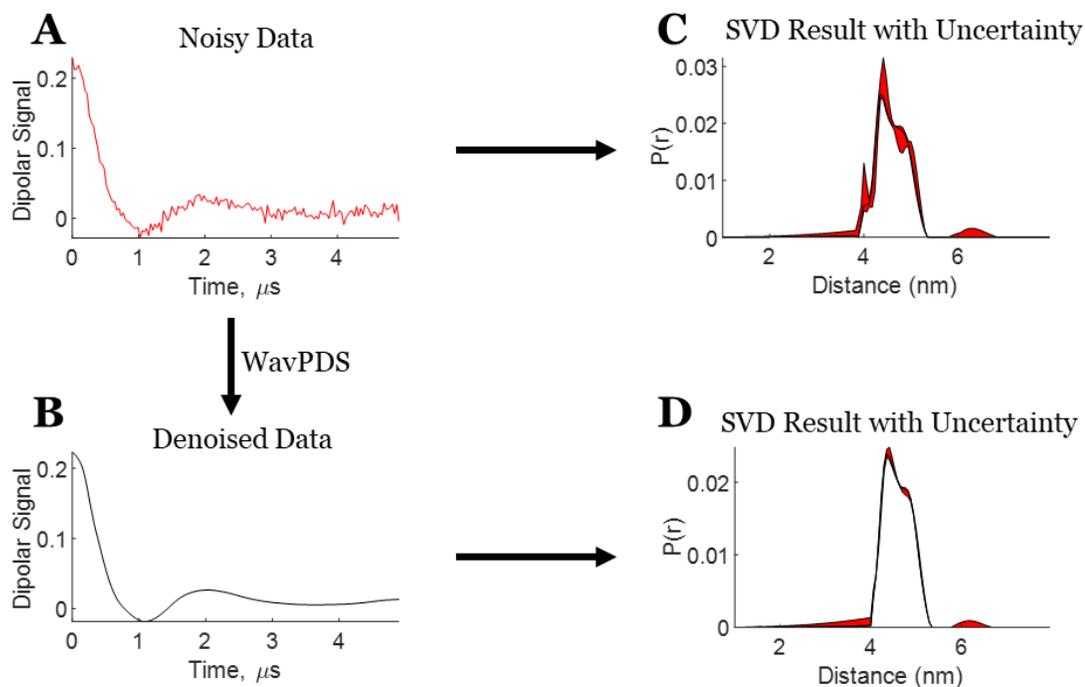
In Fig. 6.1, the step-wise process of estimating the uncertainty in a distribution was shown. The distance distribution  $P(r)$  is reconstructed from the SVD method of in equation 6.5 (cf. Fig. 6.1A). At each distance or distance range, a Piccard plot is used to obtain the SVC cut-off, and then the distribution values are plotted to obtain the regions of convergence, because the Piccard plot is in the log domain and can conceal part of converged region. The actual distribution values reveal the complete converged region. The Piccard plot in Fig. 6.1B shows that at  $2\text{ nm}$  ( $j = 19$ ) the solution after convergence remains converged till the last SVC (i.e. 155), whereas the Piccard plot at  $4.3\text{ nm}$  ( $j = 59$ , cf. Fig. 6.1C) diverges at the 57<sup>th</sup> SVC. The Piccard plot at  $2\text{ nm}$  where  $P(r) = 0$  is a special case for which the solution at that location does not diverge. Usually, the singular values that should be ideally 0, but are very close to 0 because of computer numerical round-off error, causing the solution to diverge. But, in this special case, the numerator ( $U^T_i SV_j$ ) in equation 6.6 is itself 0, preventing the divergence. The converged regions can be seen in Figs. 6.1D and 6.1E where the distribution values are plotted for  $2\text{ nm}$  (between 80 and 155 SVCs) and  $4.3\text{ nm}$  (between 13 and 56 SVCs), where values of the distribution hardly change. We refer to these as the uncertainty regions. The region outside is not converged, so incorporating their values would lead to misleading results.

The minimum and maximum values of  $P(r)$  in the converged regions is taken as the uncertainty in  $P(r)$  and is given by equations 6.7 and 6.8. This process is repeated for each distance region. Fig. 6.1F shows the distance distribution  $P(r)$  (cf. Fig. 6.1A) along with its uncertainty (in red) for each distance. It can be seen that there is very little uncertainty in the distribution, confirming the accuracy of the SVD solution.

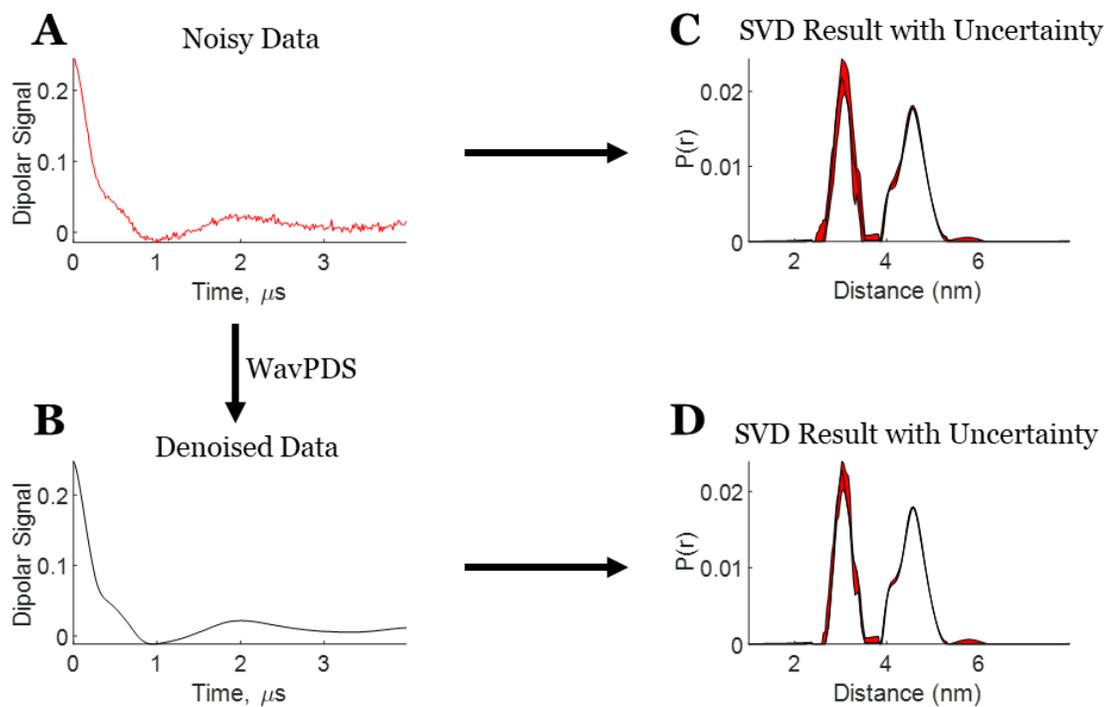
This procedure to find the uncertainty in  $P(r_j)$  for each value of  $j$  is somewhat different from the SVD method recommended in Chapter 5 to obtain the  $P(r)$  by dividing  $r$  into just three or four ranges and using a common SVC cut-off within each of these ranges.<sup>7</sup> These ranges can be selected so that within a given range of  $r$  a common cut-off exists, providing good convergence. But the extent of acceptable cut-off does vary for each  $r_j$  within a given range, thereby necessitating the more detailed analysis to provide the uncertainty.

In Cases 1 and 2 (cf. Figs. 6.2 and 6.3, respectively), the uncertainty in the distance distribution shows that the SVD solution can be obtained with high confidence for both unimodal and bimodal experimental cases. The SVD solutions for noisy and denoised signal are very similar, however, the uncertainty from the noisy data cf. (Figs. 6.2C and 6.3C) is greater than that for the denoised data (cf. Figs. 6.2D and 6.3D), as expected. Whereas high *SNR* is usually needed to obtain good results (which can be accomplished through denoising), even for *SNRs* of 38 and 80 (cf. Figs. 6.2A and Fig. 6.3A), one can obtain a reasonable distribution using the SVD approach.

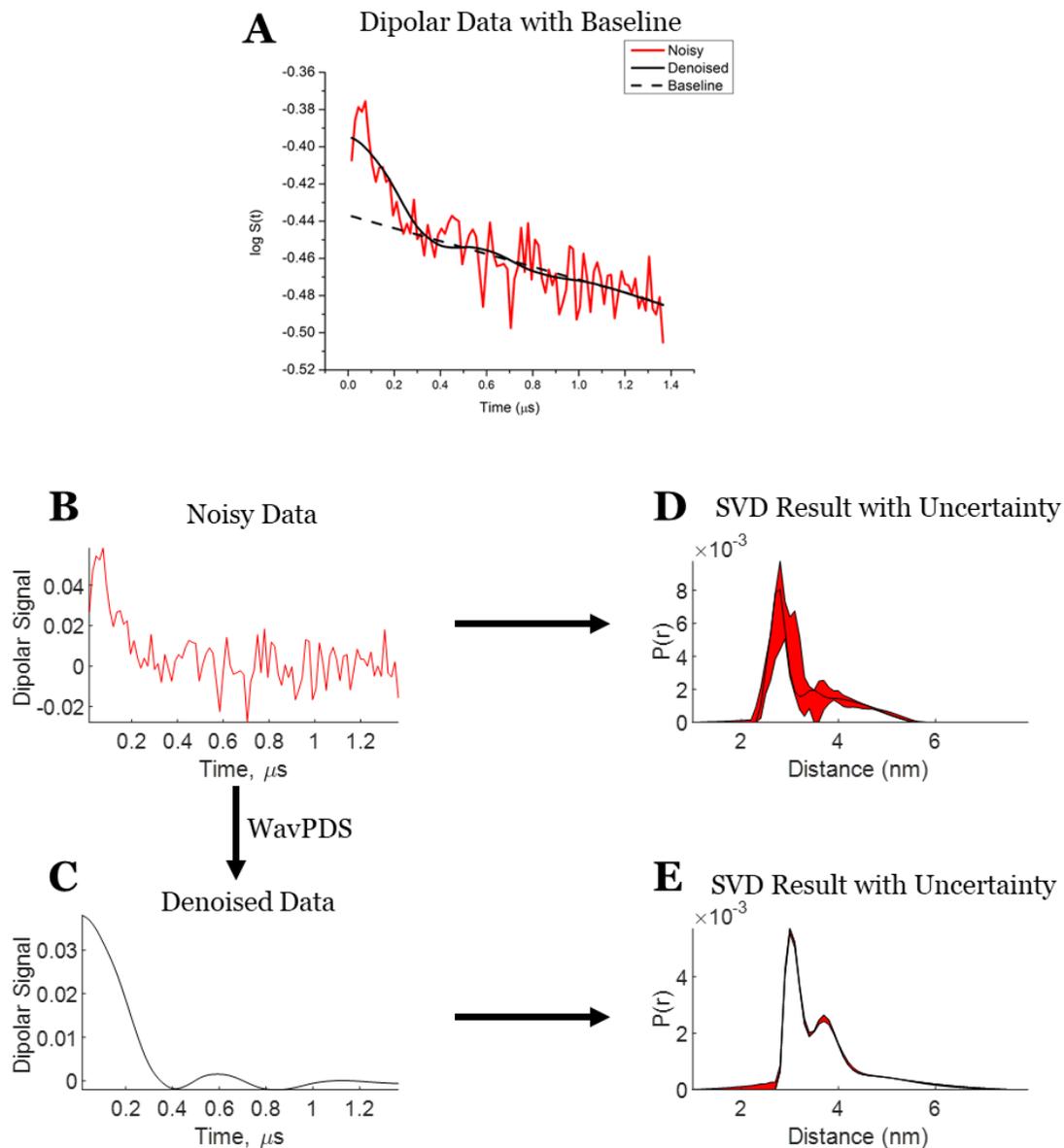
In Cases 1 and 2, the baseline subtraction was carried out on the original noisy data before applying denoising; this can negatively affect  $P(r)$  at longer distances. The uncertainty around 6 nm in Figs 6.2C and 6.3C is because of unpredictability in baseline selection in noisy data. In Case 3, the dipolar signal in Fig. 6.4 is first denoised and then the baseline is subtracted in the log-domain (with the same baseline used for both noisy and denoised data). Due to denoising, the baseline was accurately detected (cf. Fig. 6.4A) despite the poor  $SNR$  of the original data. The uncertainty analysis shows that the distance distribution of denoised data is highly reliable (cf. Figs 6.4C and 6.4E), but unlike cases 1 and 2, the SVD solution from the noisy data in case 3 is not (cf. Figs 6.4B and 6.4D). This shows that denoising is required to obtain an accurate SVD solution from noisy data.



**Figure 6.2:** Case 1 – SVD reconstruction with uncertainty analysis of the noisy and denoised dipolar signal for unimodal distribution. **A)** Noisy data; **B)** Denoised data using WavPDS; **C)** Distance distribution from noisy data with uncertainty (in red); and **D)** Distance distribution from denoised data with uncertainty (in red).



**Figure 6.3:** Case 2 – SVD reconstruction with uncertainty analysis of the noisy and denoised dipolar signal for bimodal distribution. **A)** Noisy data; **B)** Denoised data using WavPDS; **C)** Distance distribution from noisy data with uncertainty (in red); and **D)** Distance distribution from denoised data with uncertainty (in red).



**Figure 6.4:** Case 3 – SVD reconstruction with uncertainty analysis of the noisy and denoised dipolar signal at sub-micromolar concentration. **A)** Noisy (red) and denoised (black) data with baseline **B)** Noisy data; **C)** Denoised data using WavPDS; **D)** Distance distribution from noisy data with uncertainty (in red); and **E)** Distance distribution from denoised data with uncertainty (in red).

## 6.4 Conclusion

The results from the three experimental cases show that the distance distribution  $P(r)$  obtained using the SVD method is has very small uncertainty. This will be particularly important in studying cases with several overlapping and non-overlapping distributions.

In this paper, we have presented a method to determine uncertainty due to using the SVD method. This enhances the reliability of the results. Although PDS was used an example, the approach can be applied to SVD-based methods used in other spectroscopies and physical methods. It can also readily be extended to higher dimensional signals.

## References

- (1) Allen, G. I.; Peterson, C.; Vannucci, M.; Maletić-Savatić, M. Regularized Partial Least Squares with an Application to NMR Spectroscopy. *Stat. Anal. Data Min.* **2013**, *6* (4), 302–314.
- (2) Borbat, P. P.; Freed, J. H. Pulse Dipolar Electron Spin Resonance: Distance Measurements. In *Structure and Bonding*; Springer Berlin Heidelberg, 2013; pp 1–82.
- (3) Kuramshina, G. M.; Weinhold, F.; Kochikov, I. V.; Yagola, A. G.; Pentin, Y. A. Joint Treatment of Ab Initio and Experimental Data in Molecular Force Field Calculations with Tikhonov's Method of Regularization. *J. Chem. Phys.* **1994**, *100* (2), 1414–1424.
- (4) Fidêncio, P. H.; Poppi, R. J.; de Andrade, J. C. Determination of Organic Matter in Soils Using Radial Basis Function Networks and near Infrared Spectroscopy. *Anal. Chim. Acta* **2002**, *453* (1), 125–134.
- (5) Tyler, B. J.; Castner, D. G.; Ratner, B. D. Regularization: A Stable and Accurate Method for Generating Depth Profiles from Angle-Dependent XPS Data. *Surf. Interface Anal.* **1989**, *14* (8), 443–450.
- (6) Hasekamp, O. P.; Landgraf, J. Ozone Profile Retrieval from Backscattered Ultraviolet Radiances: The Inverse Problem Solved by Regularization. *J. Geophys. Res. Atmos.* **2001**, *106* (D8), 8077–8088.
- (7) Srivastava, M.; Freed, J. H. Singular Value Decomposition Method to Determine Distance Distributions in Pulsed Dipolar Electron Spin Resonance. *J. Phys. Chem. Lett.* **2017**, *8* (22), 5648–5655.
- (8) Chiang, Y.-W.; Borbat, P. P.; Freed, J. H. The Determination of Pair Distance Distributions by Pulsed ESR Using Tikhonov Regularization. *J. Magn. Reson.* **2005**, *172* (2), 279–295.
- (9) Edwards, T. H.; Stoll, S. A Bayesian Approach to Quantifying Uncertainty from Experimental Noise in DEER Spectroscopy. *J. Magn. Reson.* **2016**, *270*, 87–97.

- (10) Hustedt, E. J.; Martinelli, F.; Stein, R. A.; Faraldo-Gomez, J.; Mchaourab, H. S. Confidence Analysis of DEER Data and Its Structural Interpretation with Ensemble-Biased Metadynamics. *bioRxiv* **2018**.
- (11) Borbat, P. P.; Crepeau, R. H.; Freed, J. H. Multifrequency Two-Dimensional Fourier Transform ESR: An X/Ku-Band Spectrometer. *J. Magn. Reson.* **1997**, *127* (2), 155–167.
- (12) Polyhach, Y.; Bordignon, E.; Tschaggelar, R.; Gandra, S.; Godt, A.; Jeschke, G. High Sensitivity and Versatility of the DEER Experiment on Nitroxide Radical Pairs at Q-Band Frequencies. *Phys. Chem. Chem. Phys.* **2012**, *14* (30), 10762.
- (13) Reginsson, G. W.; Hunter, R. I.; Cruickshank, P. A. S.; Bolton, D. R.; Sigurdsson, S. T.; Smith, G. M.; Schiemann, O. W-Band PELDOR with 1kW Microwave Power: Molecular Geometry, Flexibility and Exchange Coupling. *J. Magn. Reson.* **2012**, *216*, 175–182.
- (14) Hertel, M. M.; Denysenkov, V. P.; Bennati, M.; Prisner, T. F. Pulsed 180-GHz EPR/ENDOR/PELDOR Spectroscopy. *Magn. Reson. Chem.* **2005**, *43* (S1), S248–S255.
- (15) Srivastava, M.; Anderson, C. L.; Freed, J. H. A New Wavelet Denoising Method for Selecting Decomposition Levels and Noise Thresholds. *IEEE Access* **2016**, *4*, 3862–3877.
- (16) Srivastava, M.; Georgieva, E. R.; Freed, J. H. A New Wavelet Denoising Method for Experimental Time-Domain Signals: Pulsed Dipolar Electron Spin Resonance. *J. Phys. Chem. A* **2017**, *121* (12), 2452–2465.
- (17) Borbat, P. P.; Freed, J. H. Measuring Distances by Pulsed Dipolar ESR Spectroscopy: Spin-Labeled Histidine Kinases. *Methods Enzymol.* **2007**, *423*, 52–116.
- (18) Georgieva, E. R.; Borbat, P. P.; Ginter, C.; Freed, J. H.; Boudker, O. Conformational Ensemble of the Sodium-Coupled Aspartate Transporter. *Nat. Struct. Mol. Biol.* **2013**, *20* (2), 215–221.
- (19) Jeschke, G. DEER Distance Measurements on Proteins. *Annu. Rev. Phys. Chem.* **2012**, *63* (1), 419–446.

- (20) Klare, J. P.; Steinhoff, H.-J. Structural Information from Spin-Labelled Membrane-Bound Proteins; Springer Berlin Heidelberg, 2013; pp 205–248.
- (21) Mchaourab, H. S.; Steed, P. R.; Kazmier, K. Toward the Fourth Dimension of Membrane Protein Structure: Insight into Dynamics from Spin-Labeling EPR Spectroscopy. *Structure* **2011**, *19* (11), 1549–1561.
- (22) Edwards, D. T.; Huber, T.; Hussain, S.; Stone, K. M.; Kinnebrew, M.; Kaminker, I.; Matalon, E.; Sherwin, M. S.; Goldfarb, D.; Han, S. Determining the Oligomeric Structure of Proteorhodopsin by Gd<sup>3+</sup>-Based Pulsed Dipolar Spectroscopy of Multiple Distances. *Structure* **2014**, *22* (11), 1677–1686.
- (23) DeBerg, H. A.; Bankston, J. R.; Rosenbaum, J. C.; Brzovic, P. S.; Zagotta, W. N.; Stoll, S. Structural Mechanism for the Regulation of HCN Ion Channels by the Accessory Protein TRIP8b. *Structure* **2015**, *23* (4), 734–744.
- (24) Schiemann, O.; Prisner, T. F. Long-Range Distance Determinations in Biomacromolecules by EPR Spectroscopy. *Q. Rev. Biophys.* **2007**, *40* (01), 1.
- (25) Hubbell, W. L.; Gross, A.; Langen, R.; Lietzow, M. A. Recent Advances in Site-Directed Spin Labeling of Proteins. *Curr. Opin. Struct. Biol.* **1998**, *8* (5), 649–656.
- (26) Bordignon, E. Site-Directed Spin Labeling of Membrane Proteins. In *EPR Spectroscopy*; Springer Berlin Heidelberg, 2011; pp 121–157.
- (27) Sil, D.; Lee, J. B.; Luo, D.; Holowka, D.; Baird, B. Trivalent Ligands with Rigid DNA Spacers Reveal Structural Requirements for IgE Receptor Signaling in RBL Mast Cells. *ACS Chem. Biol.* **2007**, *2* (10), 674–684.

## CHAPTER 7

### CONCLUSION

This dissertation comprises two focus areas: 1) signal denoising of experimental ESR signals using wavelets and 2) the distance reconstruction from the pulsed dipolar signals using singular value decomposition method, including calculating uncertainty in the solution.

#### 7.1 Wavelet Denoising

Wavelet transforms have been around for 25 years and are successfully implemented in image and video compression (e.g. JPEG2000 and MPEG formats), among other applications. It has also been used to remove noise from signal but has failed to achieve its true potential in the past. Despite the fact that noise and signal can be distinguished and separated in the wavelet domain, *standard* wavelet denoising methods are unable to calculate objective parameters to achieve that. We developed an objective method that is able to successfully identify and separate noise from signal in wavelet domain. This has led the removal of noise in ESR signals at and above SNR of 3. Due to this, the signal averaging times has been reduced by two orders of magnitude. The dissertation showed the step-by-step process that leads to denoised pulsed and cw-ESR signal: 1) representation and analysis of signal in the wavelet domain, in particular the observation of a signal at different frequency sub-bands and 2) identification and elimination of noise coefficients from signal coefficients in wavelet domain. In the denoising process, we demonstrated 1) a method to select the

frequency sub-bands dominated by noise, 2) a new formula to calculate noise thresholds that does not require noise estimation, 3) the rationale and process of using separate noise thresholds for positive and negative wavelet coefficients, 4) implementation of noise thresholding at low frequencies, and 5) flexibility to adjust the noise thresholds to remove systemic noise.

The application of wavelet denoising method in ESR has potential to carry out biological studies at low concentrations and native environments as well as for unstable biomolecules.

## **7.2 Distance Reconstruction in Pulsed Dipolar Spectroscopy**

Pulsed Dipolar Spectroscopy (PDS) methods, such as Double Electron Electron Resonance and Double Quantum Coherence, are powerful methods for studying the structure and function of biological systems. In PDS, a dipolar signal is acquired from the interaction between a pair of spin labels, from which the distance distribution between them,  $P(r)$  may be obtained between the distance ranges of 1 to 10 nm. However, due to the ill-posed nature of the inversion of the dipolar signal to yield the  $P(r)$ , one must resort to regularization or model fitting methods to obtain reasonable results. The method of Tikhonov regularization (TIKR) is commonly used, but it relies heavily on the choice of regularization parameter that yields a compromise between good resolution and stability of the  $P(r)$ . Model fitting methods, on the other hand, require *a priori* model functions to estimate  $P(r)$ , which may not accurately represent

the actual distance distributions. This is especially true if the  $P(r)$  is multimodal. We developed a new and objective approach based on singular value decomposition (SVD) that yields an optimum approximate solution, obviating the need for regularization. Instead of solving for the complete distance distribution all at once, the method finds the optimal distribution value at each distance or distance range by determining each of their different singular value cut-offs. The new method ensures optimal convergence at all distance ranges, while preventing a premature or unstable solution at some or all distance ranges. We tested the new SVD method on several model and experimental dipolar signals with unimodal and multimodal distributions. The method yields high resolution  $P(r)$  without any spurious peaks or negative  $P(r)$ 's and consistently performs better than TIKR. The new method can successfully reconstruct multimodal distributions, both overlapping and independent, with varying distribution widths.

# APPENDIX 1

## WAVELET FILTERS: FILTER VALUES AND MATLAB SCRIPT

### AX.1 Filter Values

#### AX.1.1 Coiflet-3 (coif3) Wavelet

Index	Low pass: FWT	High pass: FWT	Low pass: IFFT	High pass: IFFT
0	-3.46E-05	0.003794	-0.00379	-3.46E-05
1	-7.10E-05	0.007783	0.007783	7.10E-05
2	0.000466	-0.02345	0.023453	0.000466
3	0.001118	-0.06577	-0.06577	-0.00112
4	-0.00257	0.061123	-0.06112	-0.00257
5	-0.00901	0.405177	0.405177	0.009008
6	0.015881	-0.79378	0.793777	0.015881
7	0.034555	0.428483	0.428483	-0.03456
8	-0.0823	0.0718	-0.0718	-0.0823
9	-0.0718	-0.0823	-0.0823	0.0718
10	0.428483	-0.03456	0.034555	0.428483
11	0.793777	0.015881	0.015881	-0.79378
12	0.405177	0.009008	-0.00901	0.405177
13	-0.06112	-0.00257	-0.00257	0.061123
14	-0.06577	-0.00112	0.001118	-0.06577
15	0.023453	0.000466	0.000466	-0.02345
16	0.007783	7.10E-05	-7.10E-05	0.007783
17	-0.00379	-3.46E-05	-3.46E-05	0.003794

### Ax.1.2 Daubechies-6 (db6) Wavelet

<b>Index</b>	<b>Low pass: FWT</b>	<b>High pass: FWT</b>	<b>Low pass: IFFT</b>	<b>High pass: IFFT</b>
0	-1.08E-03	-0.11154	0.111541	-1.08E-03
1	4.78E-03	0.494624	0.494624	-4.78E-03
2	0.000554	-0.75113	0.751134	0.000554
3	-0.03158	0.31525	0.31525	0.031582
4	0.027523	0.226265	-0.22626	0.027523
5	0.097502	-0.12977	-0.12977	-0.0975
6	-0.12977	-0.0975	0.097502	-0.12977
7	-0.22626	0.027523	0.027523	0.226265
8	0.31525	0.031582	-0.03158	0.31525
9	0.751134	0.000554	0.000554	-0.75113
10	0.494624	-0.00478	0.004777	0.494624
11	0.111541	-0.00108	-0.00108	-0.11154

## AX.2 MATLAB Script to Generate Filters

```
clc;
clear all;
close all;

% Assigning wavelet
wname = 'db6'; % for Daubechies-6
wname = 'coif3'; % for Coiflet-3

% in the above lines, comment the wavelet that is not used.

% Generating filters using built-in function
[Lo_FWT,Hi_FWT,Lo_IFWT,Hi_IFWT] = wfilters(wname);

% Plotting the filters

figure,
subplot(221); stem(Lo_FWT,'LineWidth',3);
title('DWT/FWT low pass filter');
set(gca,'FontSize',30)

subplot(222); stem(Hi_FWT,'LineWidth',3);
title('DWT/FWT high pass filter');
set(gca,'FontSize',30)

subplot(223); stem(Lo_IFWT,'LineWidth',3);
title('IDWT/IFWT low pass filter');
set(gca,'FontSize',30)

subplot(224); stem(Hi_IFWT,'LineWidth',3);
title('IDWT/IFWT high pass filter');
set(gca,'FontSize',30)
```