

Monday, October 23, 2017

SUPPLEMENTARY MATERIAL

Singular value decomposition based method to determine distance distributions in pulsed dipolar spectroscopy.

Madhur Srivastava^{1,2}, and Jack H. Freed^{1,3}*

¹National Biomedical Center for Advanced ESR Technology, Cornell University, Ithaca, NY 14853, USA

²Meinig School of Biomedical Engineering, Cornell University, Ithaca, NY 14853, USA

³Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853, USA

*Corresponding Author: jhf3@cornell.edu

Monday, October 23, 2017

TABLE OF CONTENTS

S1. Details of Platform and Software

S2. Theorem for Obtaining Exact Solution using SVD

S3. Tikhonov Regularization, its SVD Form, and its drawback with respect to λ

S4. Orthogonality Validation based on Theorem for Unimodal and Bimodal Model Data

S5. Reconstructed Unimodal and Bimodal Distance Distributions of Noiseless and Some Noise Model Data with Different Singular Value Cut-off

S6. Reconstructed Bimodal Distance Distributions of Noiseless and Some Noise Model Data using new SVD method

S7. Comparison of Distance Distributions Reconstructed from TIKR, TIKR+MEM and SVD Method for Unimodal and Bimodal Dipolar Signal

S8. Piccard Plots for Individual r -values at Each Segmented Region

S1. 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)² and Maximum Entropy Method (MEM)³ codes written for MATLAB available on ACERT website (https://acert.cornell.edu/index_files/acert_resources.php) were used to generate $P(r)$.

S2. 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 .⁴ To illustrate, in SVD form, $KP = S$ can be written as (cf. eq. 3),

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

where U is an $M \times M$ column-orthogonal matrix, V is an $N \times N$ column-orthogonal matrix, and Σ is an $M \times N$ diagonal matrix containing the non-negative singular values (σ) in decreasing order. To solve for P , equation S1 can be rewritten as (cf. eq. 4),

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

or,

$$P = \sum_{i=1}^N \frac{u_i^T S}{\sigma_i} v_i \tag{S3}$$

or,

Monday, October 23, 2017

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

where u_i and v_i are the column vectors of the matrices U and V , respectively, σ_i are the individual singular values of the diagonal matrix Σ .

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,

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

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

S3. TIKHONOV REGULARIZATION, ITS SVD FORM AND ITS DRAWBACK WITH RESPECT TO λ

Tikhonov regularization (TIKR) minimizes the following function, ^{2,3}

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

where λ is the regularization parameter and L is a differential operator. The solution P_λ can be obtained as,

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

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

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

P_λ in equation S8 can be rewritten as a modified sum of singular value contributions (SVCs) as,

Monday, October 23, 2017

$$P_\lambda = \sum_{i=1}^{\sigma_i \geq \lambda} f_i \frac{u_i^T S}{\sigma_i} v_i, \quad \text{where } f_i \equiv \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \quad (\text{S9})$$

where f_i is a filter function that suppresses the SVCs from small singular values such that $\sigma_i < \lambda$. The choice of λ is crucial; one such approach is to obtain it by the L -curve method.^{5,6} Scheme 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 $q = i^{\text{th}}$ 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 λ . The method is designed to solve least squares involving the whole distribution.

REFERENCES

- (1) 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.
- (2) 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.
- (3) 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.
- (4) Horn, R. A.; Johnson, C. A. *Matrix Analysis*, First Edit.; Cambridge University Press: New York, 1985.
- (5) Hansen, P. C. Analysis of Discrete Ill-Posed Problems by Means of the L-Curve. *SIAM Rev.* **1992**, *34* (4), 561–580.
- (6) 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.

S4. Orthogonality Validation based on Theorem for Unimodal and Bimodal

Model Data: Figs. S1, S2

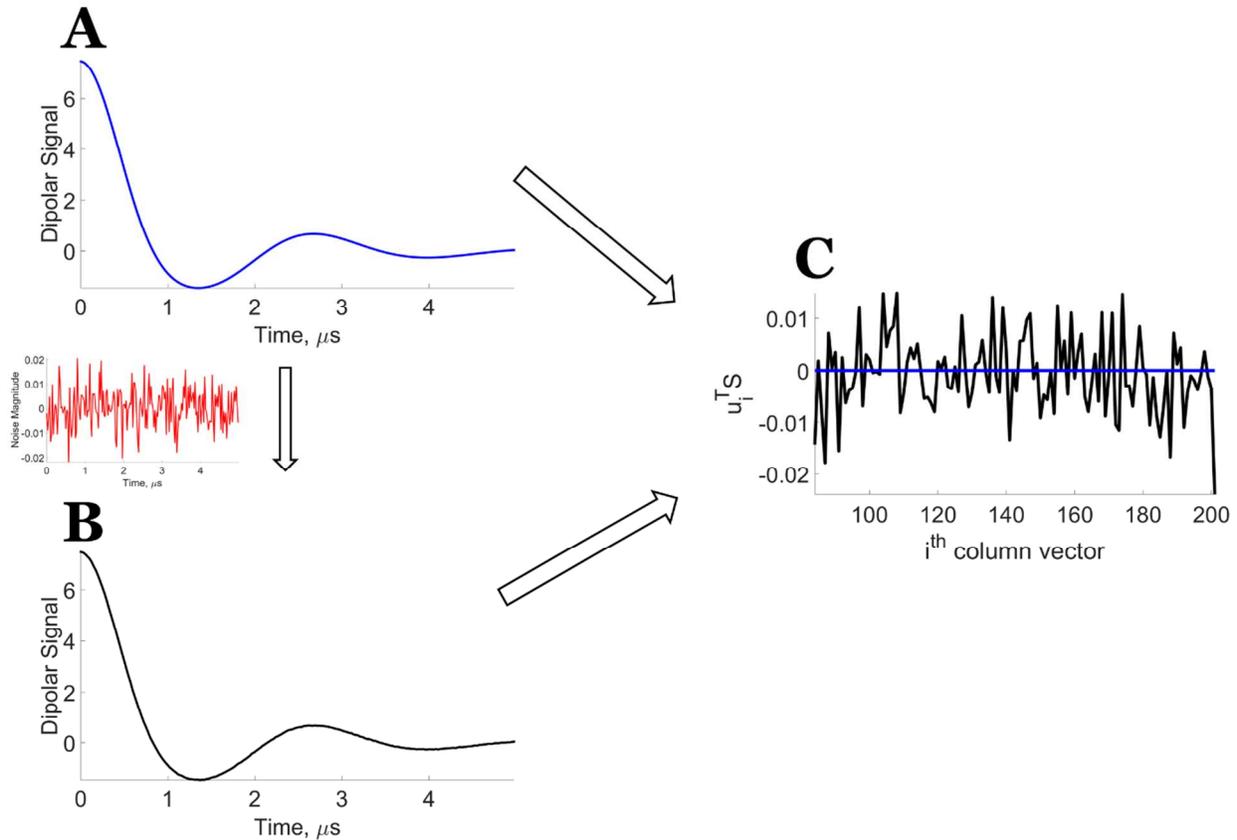


Figure S1: Unimodal Model Data -- Orthogonality validation for the noiseless data and data with some noise ($SNR \approx 850$). **A)** Noiseless data; **B)** Data with some noise; and **C)** Comparison of results of $u_i^T S$ at i^{th} column vector for noiseless data and data with some noise. Signal 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.

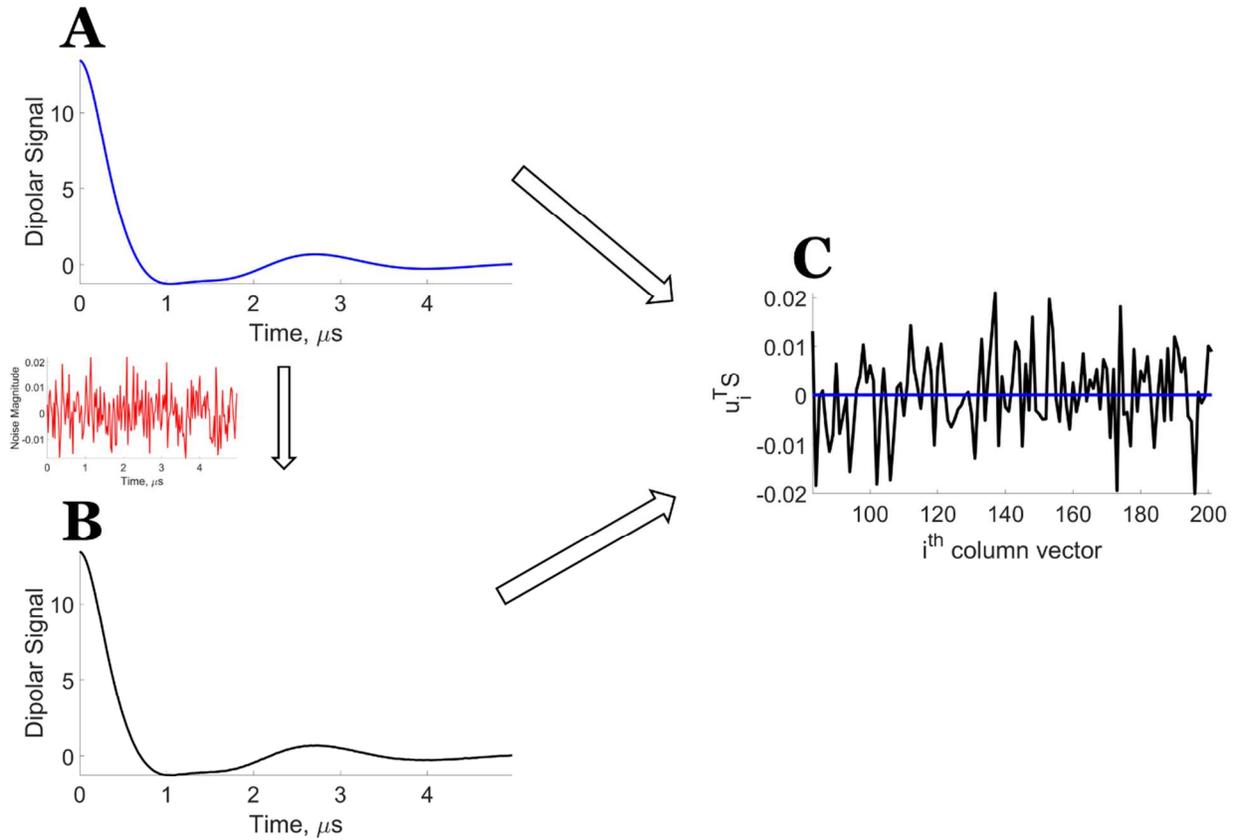
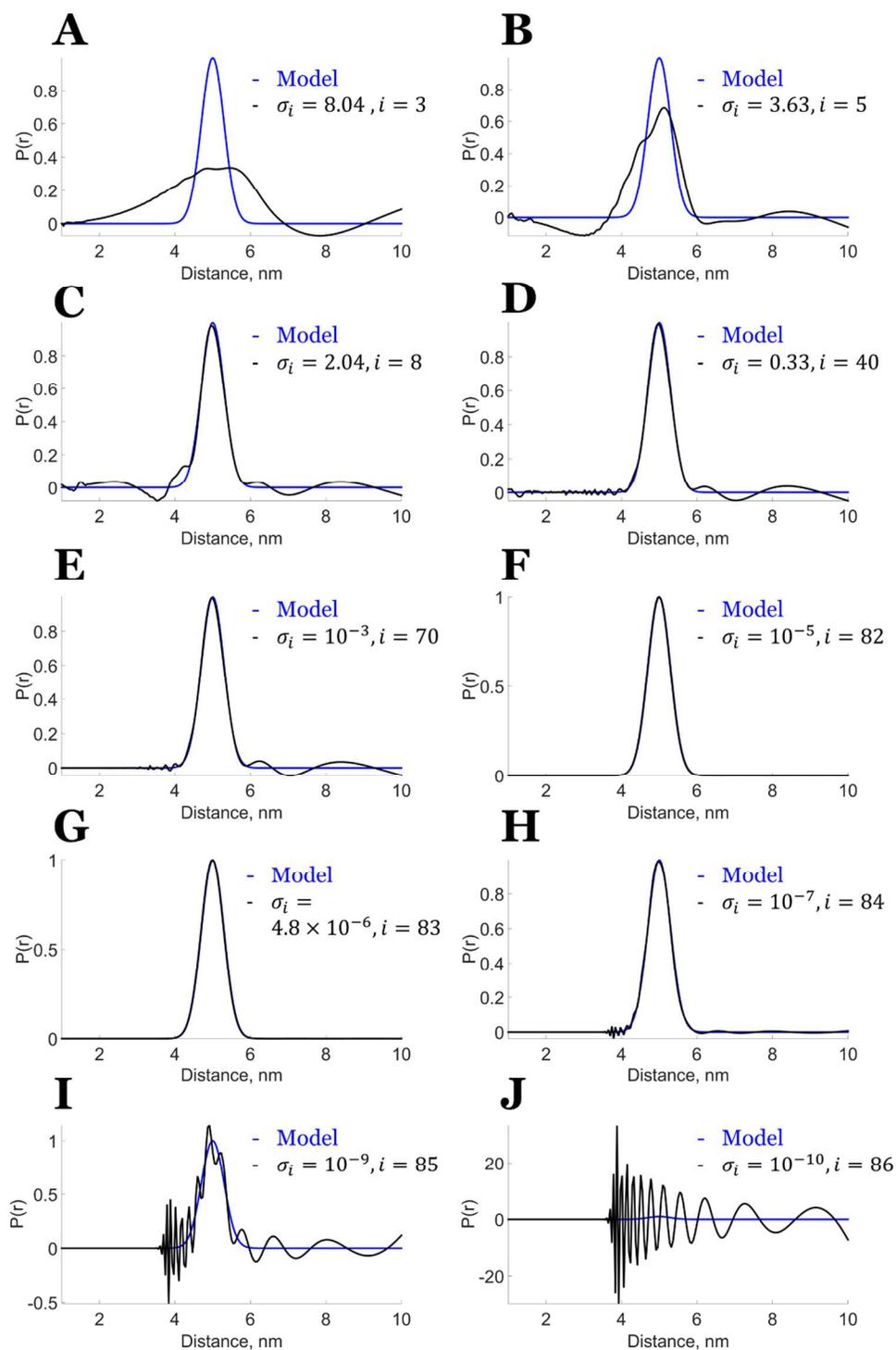


Figure S2: Bimodal Model Data -- Orthogonality validation for the noiseless data and data with some noise ($SNR \approx 850$). **A)** Noiseless data; **B)** Data with some noise; and **C)** Comparison of results of $u_i^T S$ at i^{th} column vector for noiseless data and data with some noise. Signal 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.

S5. Reconstructed Unimodal and Bimodal Distance Distributions of Noiseless and Some Noise Model Data with Different Singular Value Cut-offs: Figs. S4-S7



Monday, October 23, 2017

Figure 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)** $i = 86$, $\sigma_i = 10^{-10}$.

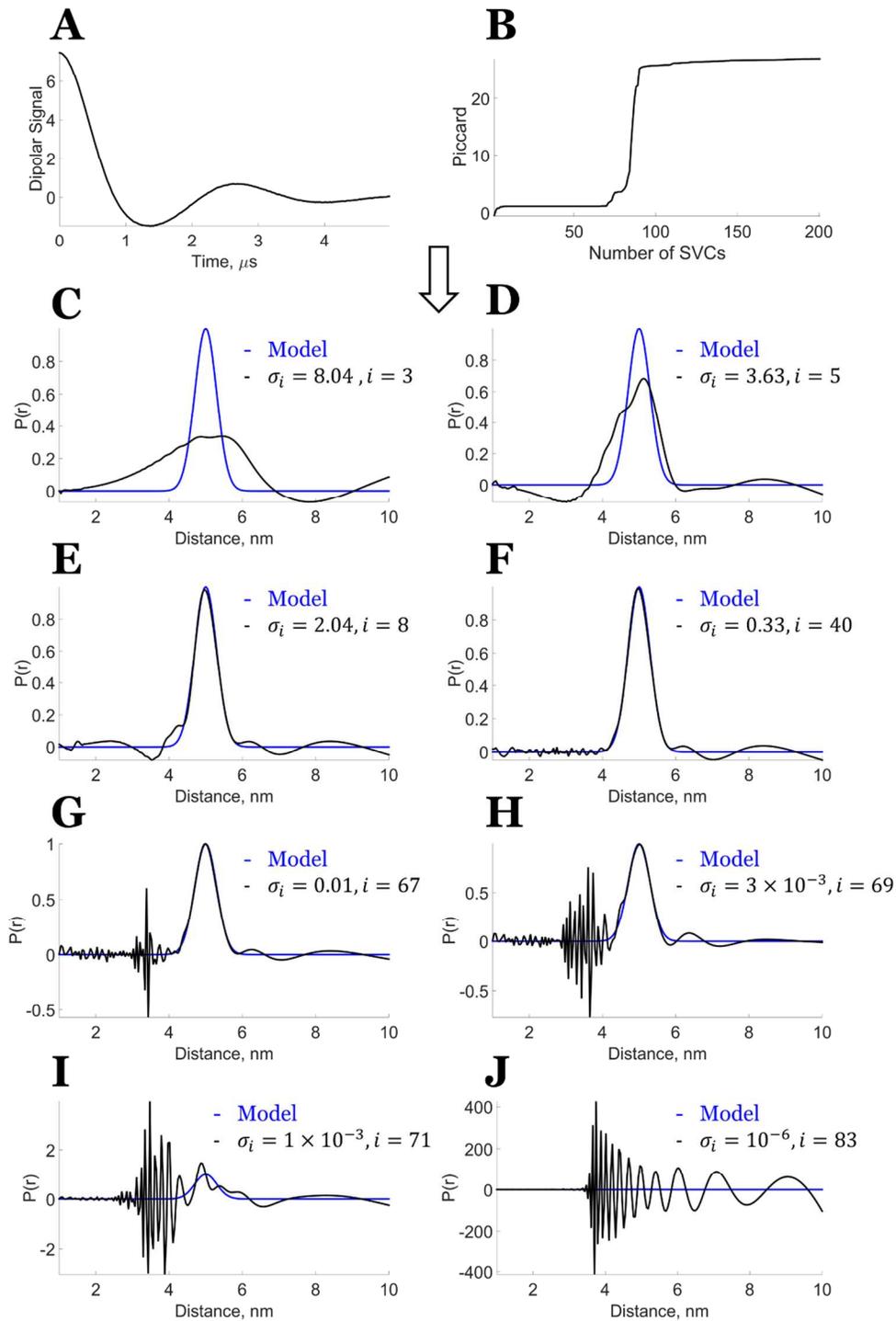


Figure S5: Model Data with Some Noise ($SNR \approx 850$) -- Unimodal Distance Distribution. **A)** Noisy model data dipolar signal; **B)** Piccard plot of the unimodal distribution from the noisy model data at different number of Singular Value

Monday, October 23, 2017

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**) $i = 83$, $\sigma_i = 10^{-6}$.

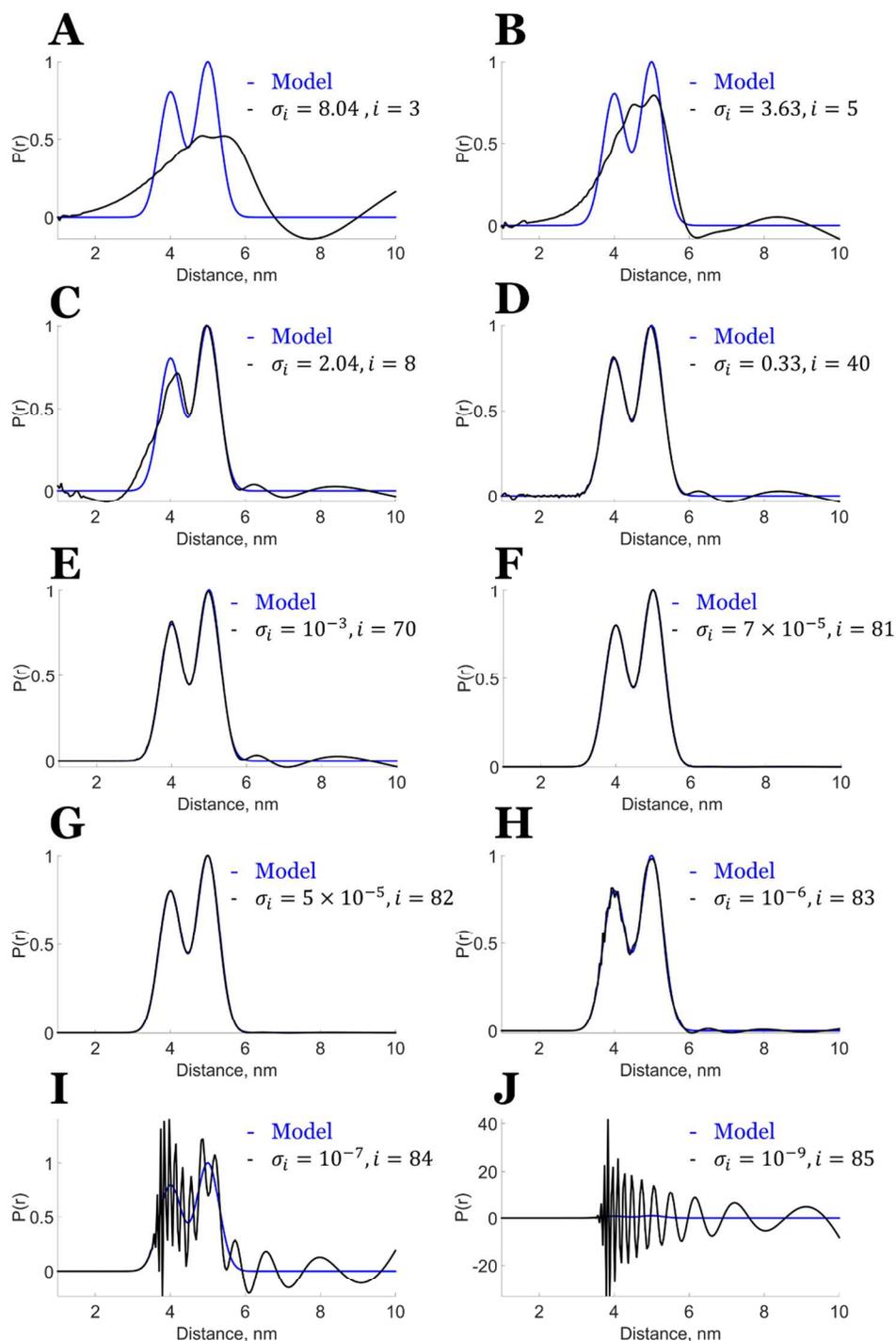


Figure S6: Model data noise-free case -- Bimodal Distance Distribution. Examples of $P(r)$ obtained for different values of the singular value cut-off. **A)** $i = 3, \sigma_i = 8.04$; **B)** $i = 5, \sigma_i = 3.63$; **C)** $i = 8, \sigma_i = 2.04$; **D)** $i = 40, \sigma_i = 0.33$; **E)** $i = 70, \sigma_i = 10^{-3}$; **F)** $i =$

Monday, October 23, 2017

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**) $i = 85$, $\sigma_i = 10^{-9}$.

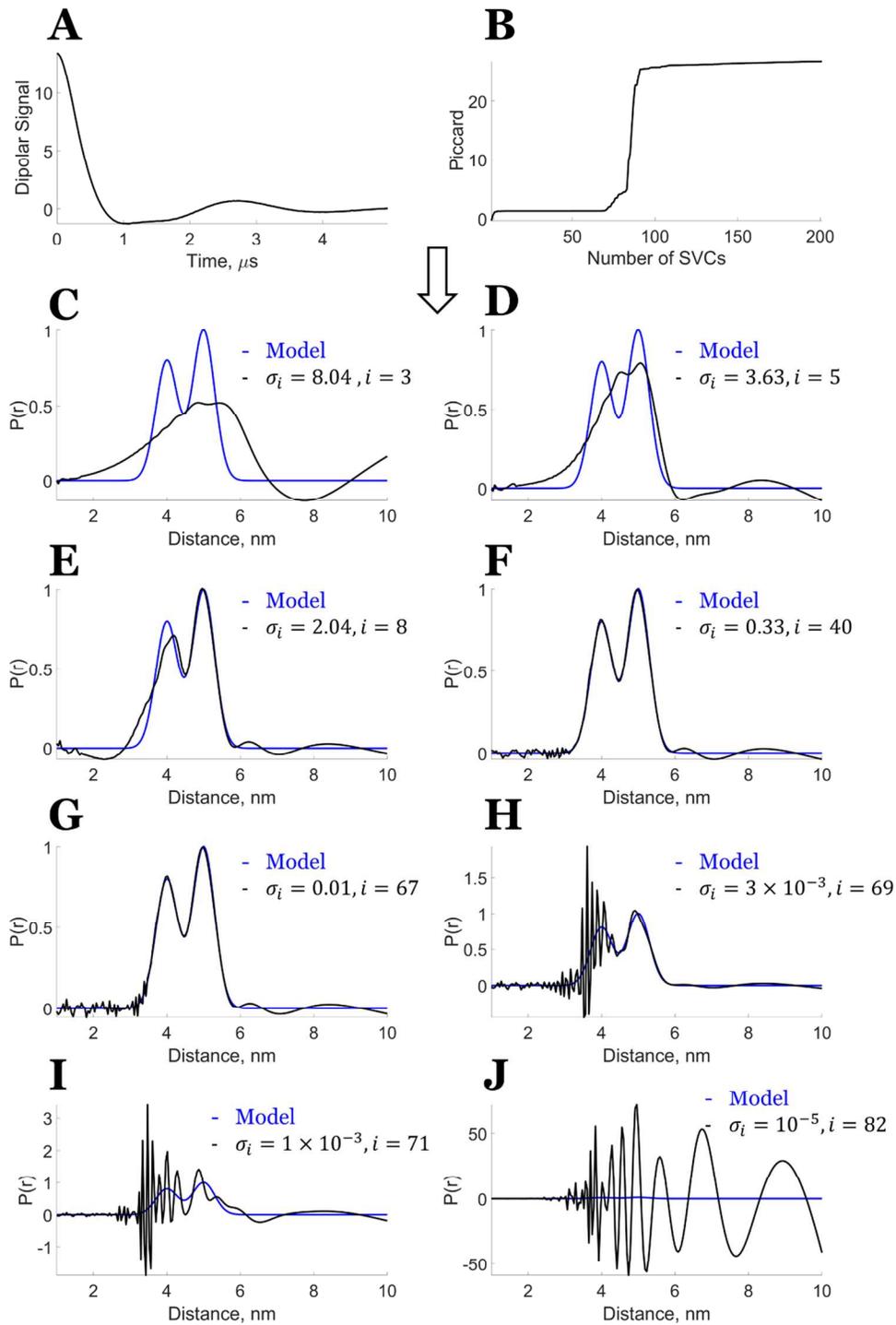


Figure S7: Model Data with Some Noise ($SNR \approx 850$) -- Unimodal Distance Distribution. **A)** Noisy model data dipolar signal; **B)** Piccard plot of the unimodal distribution from the noisy model data at different number of Singular Value

Monday, October 23, 2017

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**) $i = 82$, $\sigma_i = 10^{-5}$.

S6. Reconstructed Bimodal Distance Distributions of Noiseless and Some Noise Model Data using new SVD method

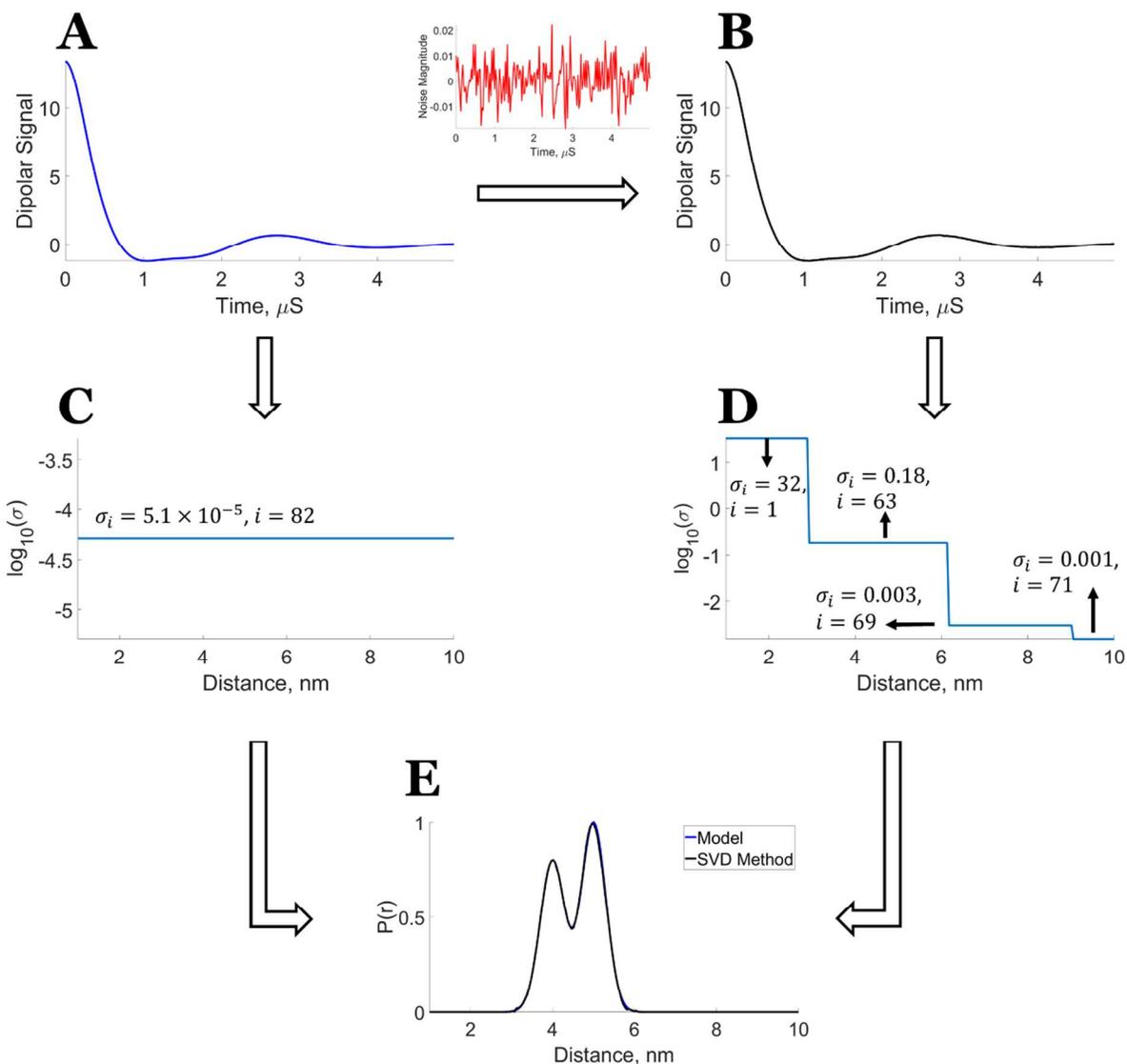


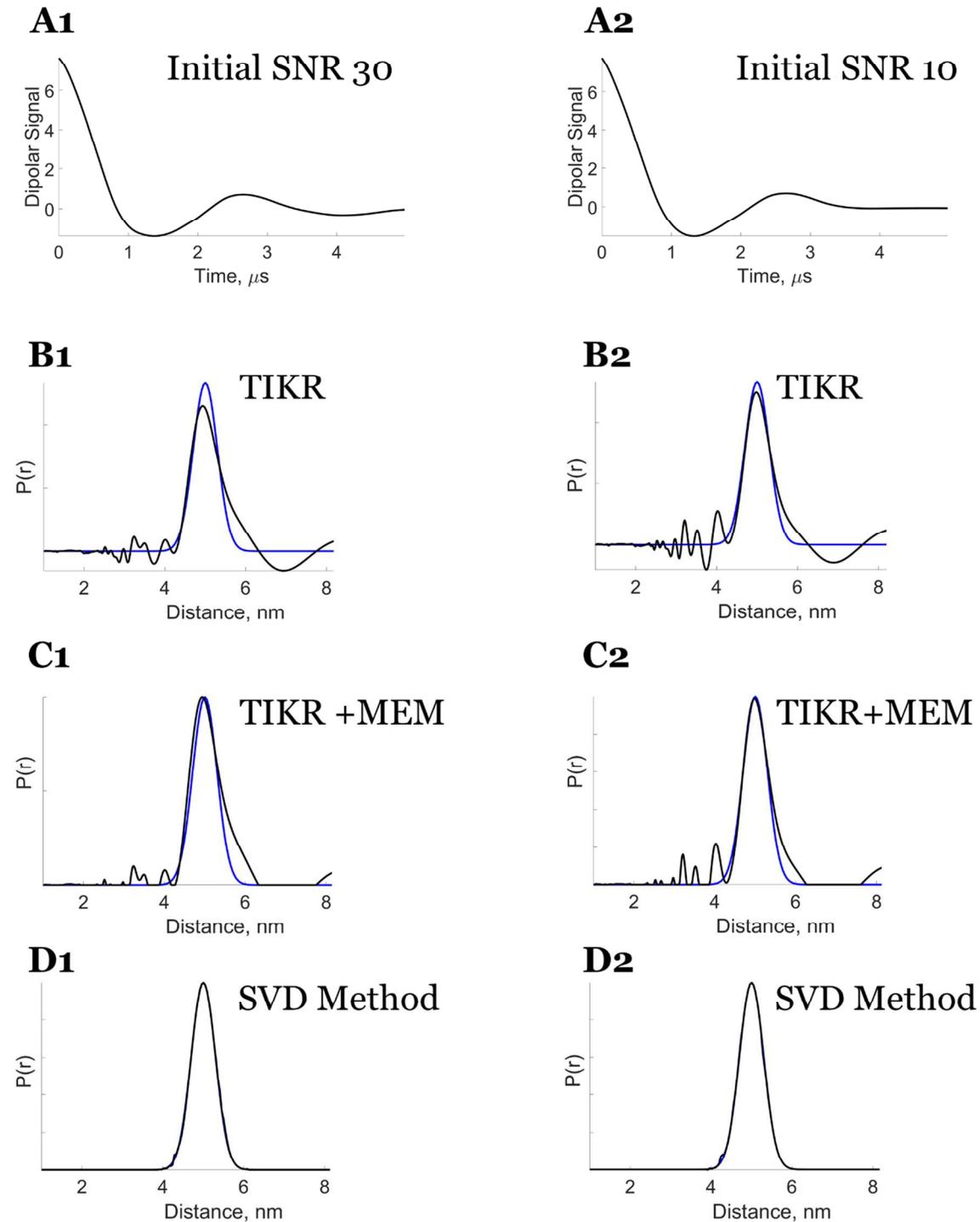
Figure S8: Bimodal Model -- Reconstruction of distance distribution for noise-free model data and noisy model data ($SNR \approx 850$) using the new SVD method. **A)** Model dipolar signal; **B)** Model dipolar signal with added noise (see added noise in Red plot); **C)** Singular value cut-off at each 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

Monday, October 23, 2017

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.

S7. Comparison of Distance Distributions Reconstructed from TIKR, TIKR+MEM and SVD Method for Unimodal and Bimodal Dipolar Signal:

Figs. S9-S12 and Table S1



Monday, October 23, 2017

Figure 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).

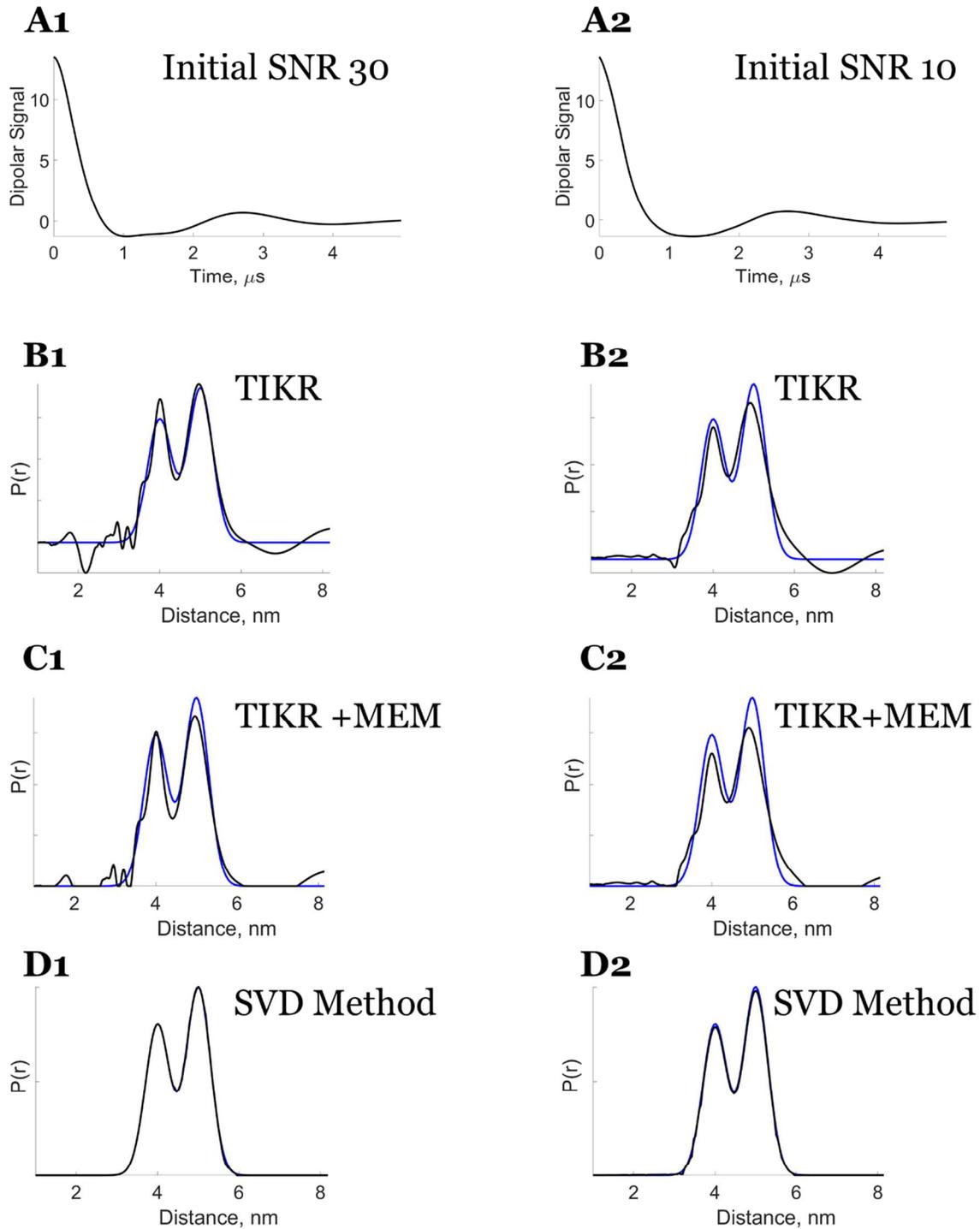


Figure S10: Model Data -- Bimodal: Comparison of distance distributions reconstructed from Tikhonov regularization method (TIKR), Tikhonov regularization followed by maximum entropy method (TIKR+MEM), and new SVD method for model

Monday, October 23, 2017

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).

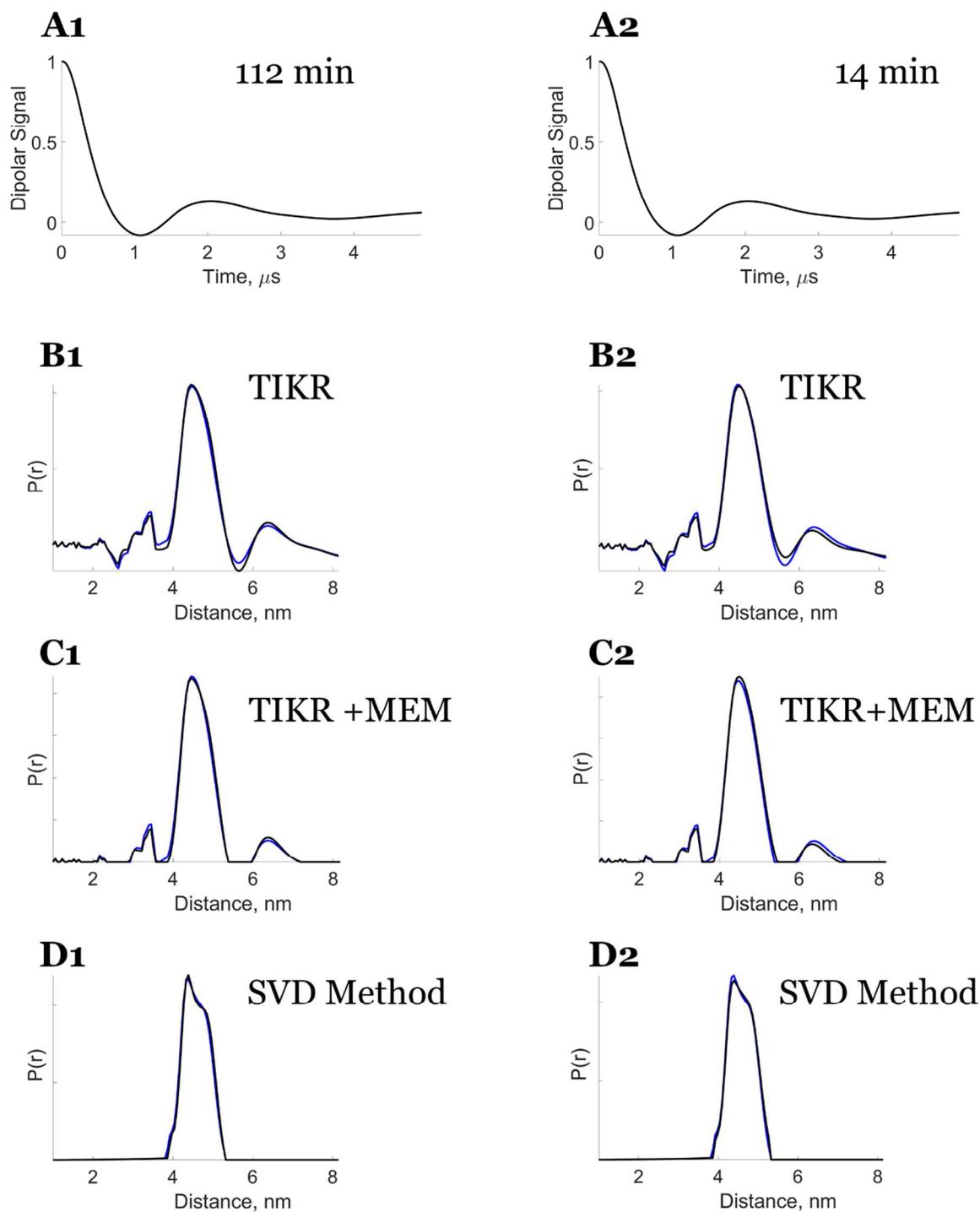


Figure S11: Experimental Data -- Unimodal: Comparison of distance distributions reconstructed from Tikhonov regularization method (TIKR), Tikhonov regularization followed by maximum entropy method (TIKR+MEM), and new SVD method for

Monday, October 23, 2017

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.¹

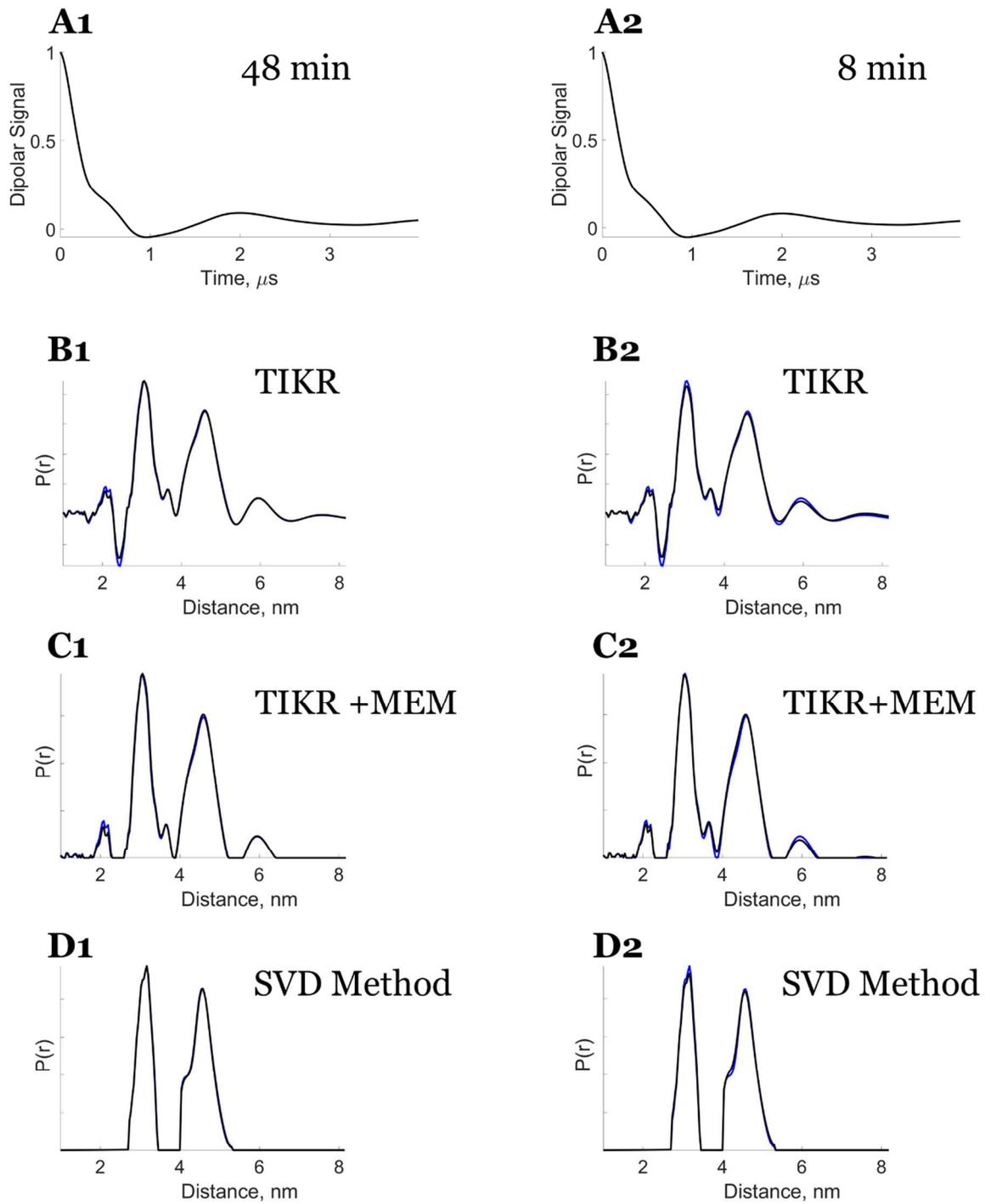


Figure S12: Experimental Data -- Bimodal: Comparison of distance distributions reconstructed from Tikhonov regularization method (TIKR), Tikhonov regularization followed by maximum entropy method (TIKR+MEM), and new SVD method for

Monday, October 23, 2017

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

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

S8. Piccard Plots for Individual r -values at Each Segmented Region: Figs. S13-S20

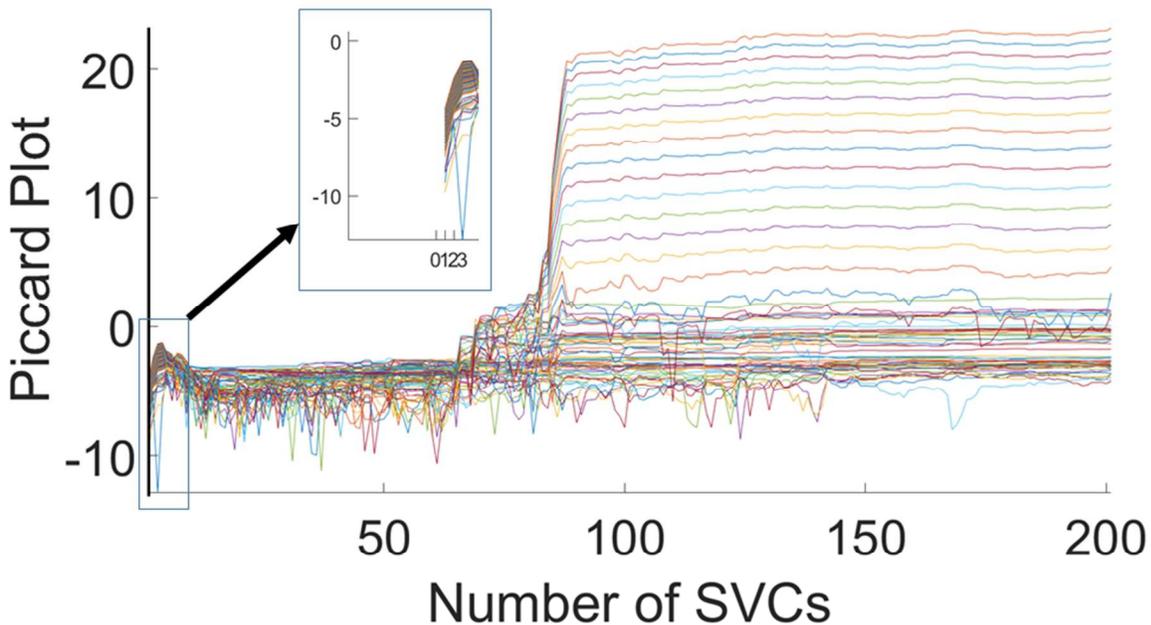
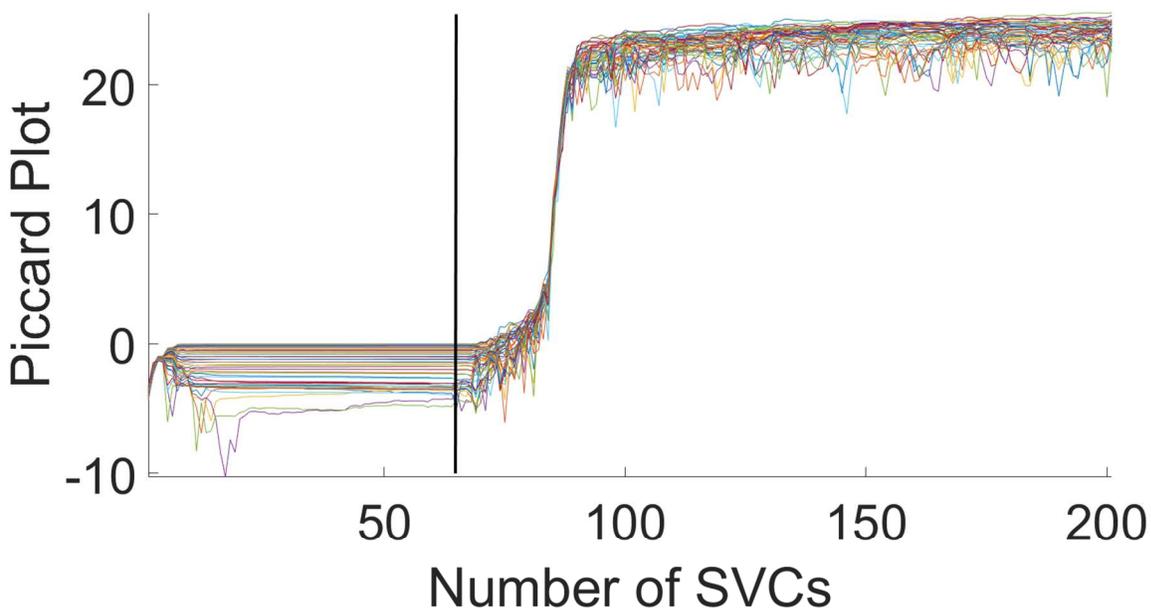


Figure S13: Model Data – Unimodal: Singular value cut-off on the Piccard plot for segment 1 (1 to 3.88 nm) of r region.



Monday, October 23, 2017

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

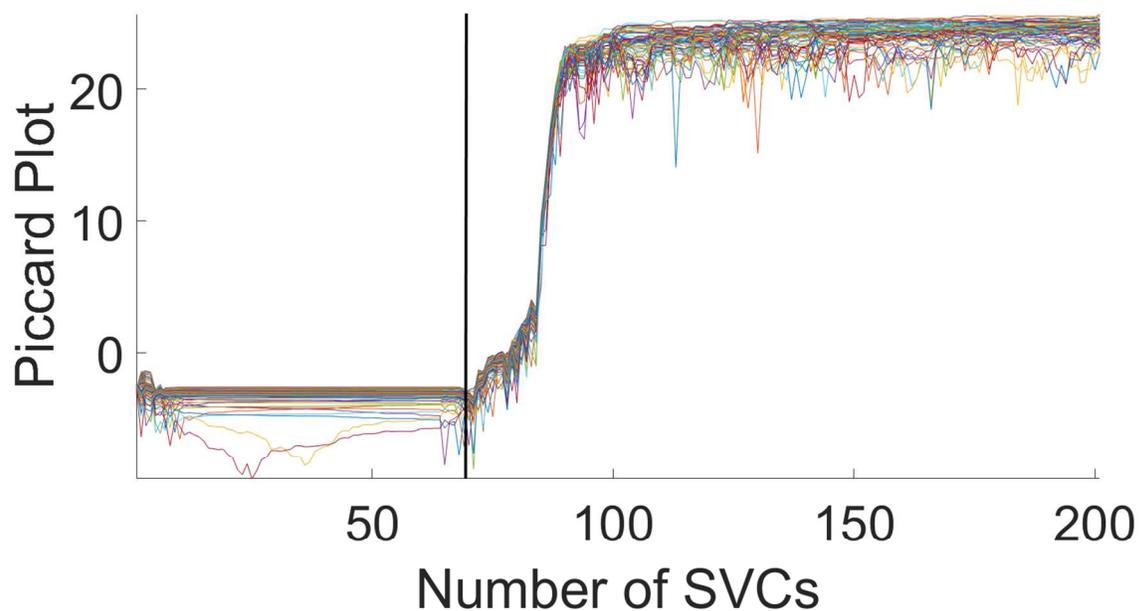


Figure S15: Model Data – Unimodal: Singular value cut-off on the Piccard plot for segment 2 (6.13 to 8.96 nm) of r region.

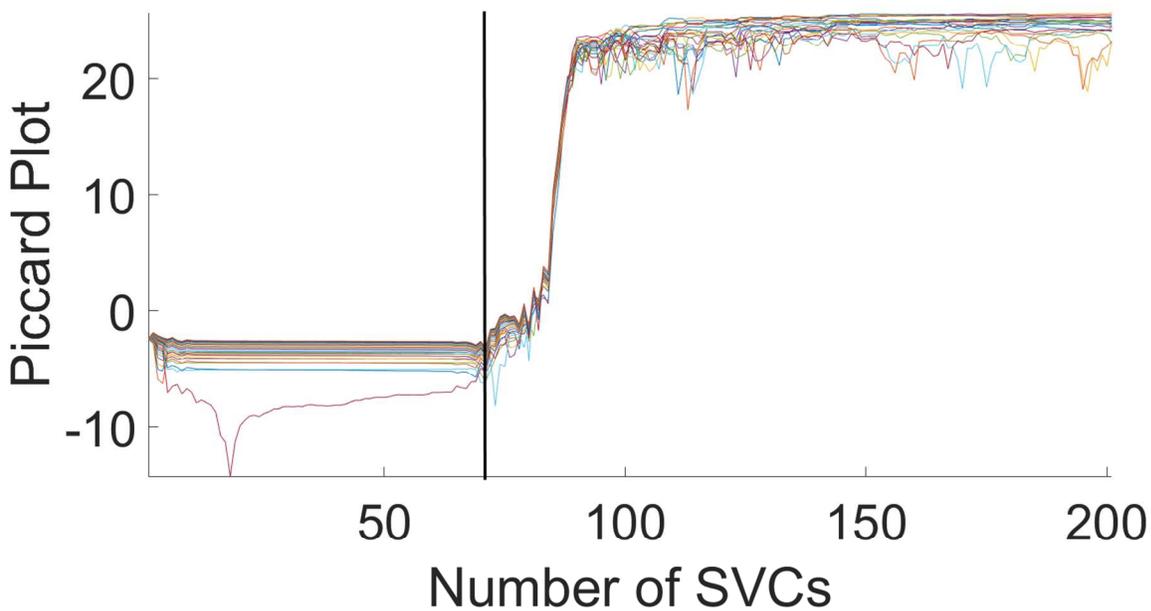


Figure S16: Model Data – Unimodal: Singular value cut-off on the Piccard plot for segment 2 (9.01 to 10 nm) of r region.

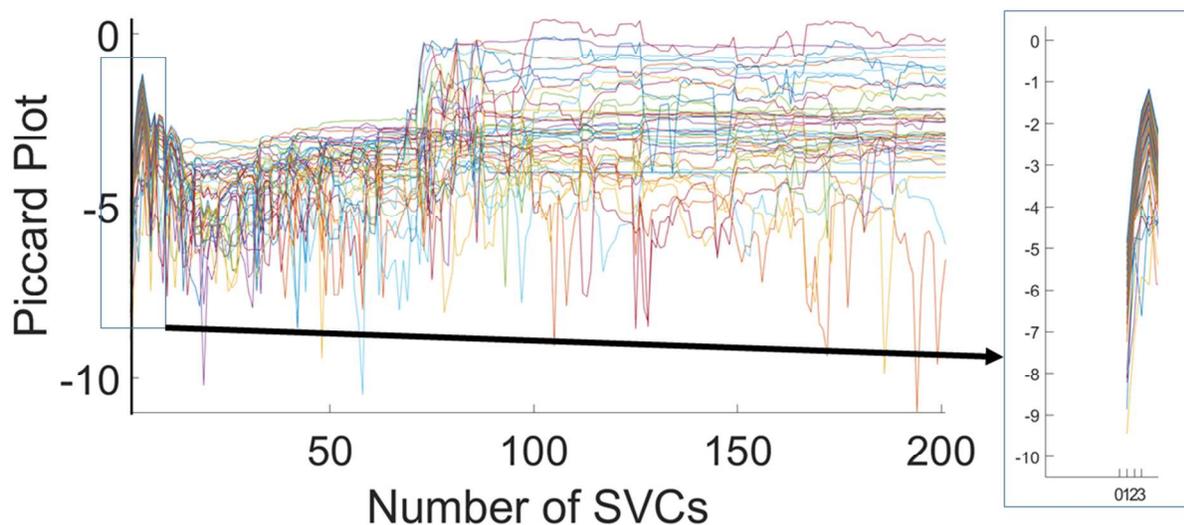


Figure S17: Model Data – Bimodal: Singular value cut-off on the Piccard plot for segment 1 (1 to 2.89 nm) of r region.

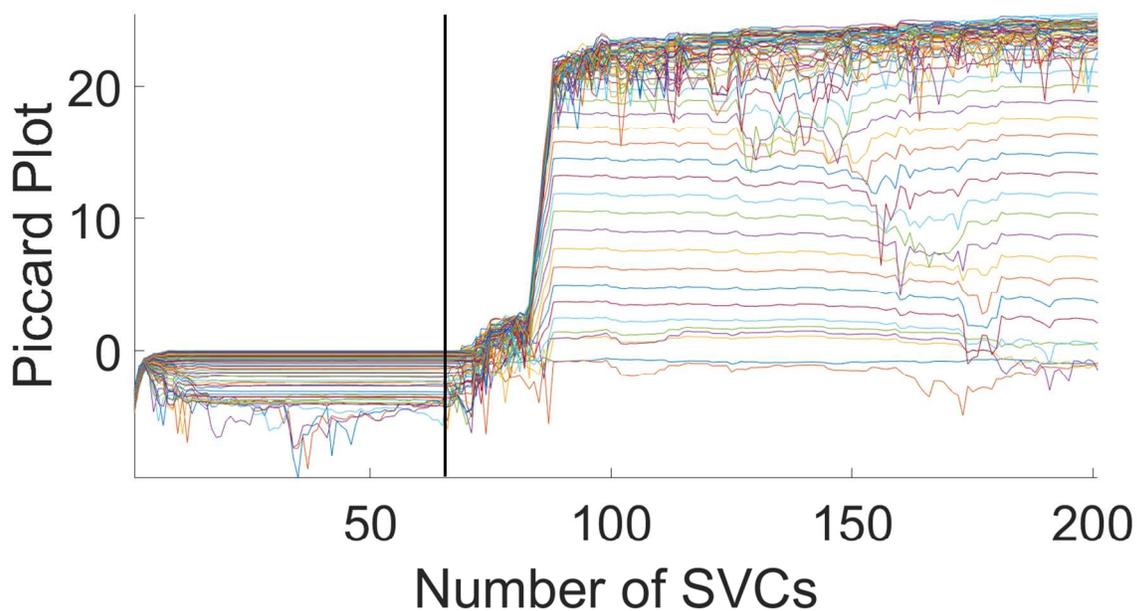


Figure S18: Model Data – Bimodal: Singular value cut-off on the Piccard plot for segment 2 (2.93 to 6.08 nm) of r region.

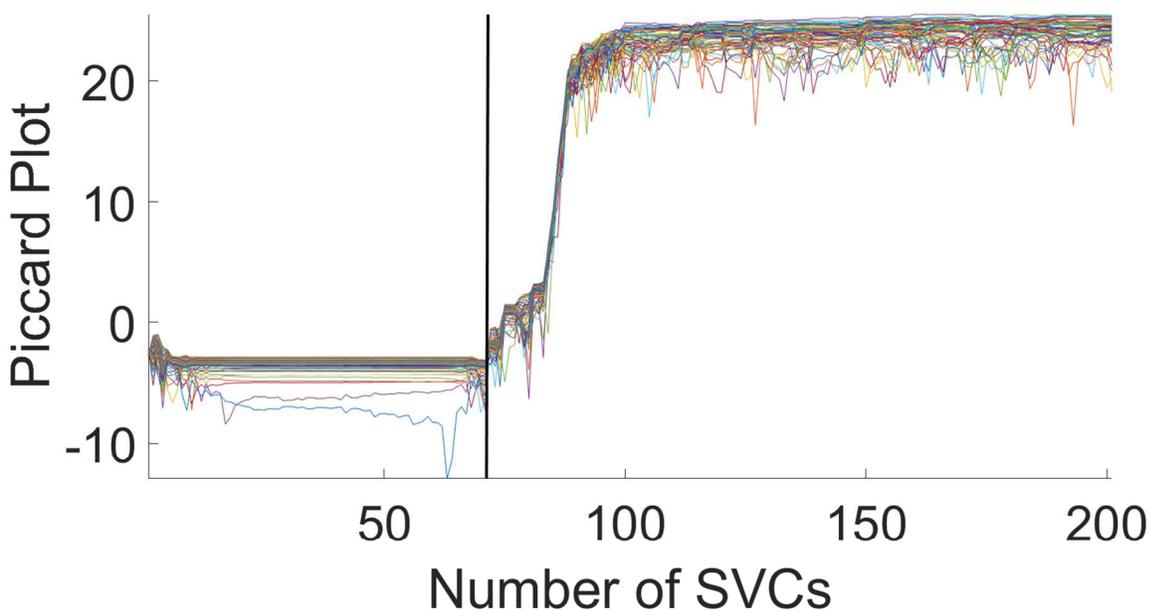


Figure S19: Model Data – Bimodal: Singular value cut-off on the Piccard plot for segment 3 (6.13 to 8.96 nm) of r region.

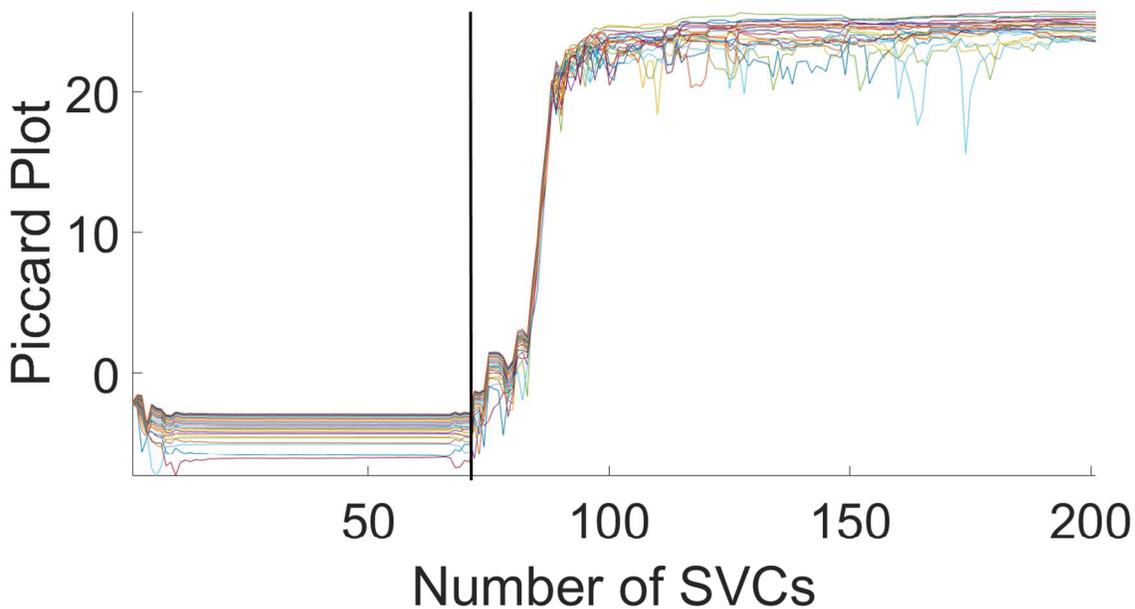


Figure S20: Model Data – Bimodal: Singular value cut-off on the Piccard plot for segment 4 (9.01 to 10 nm) of r region.

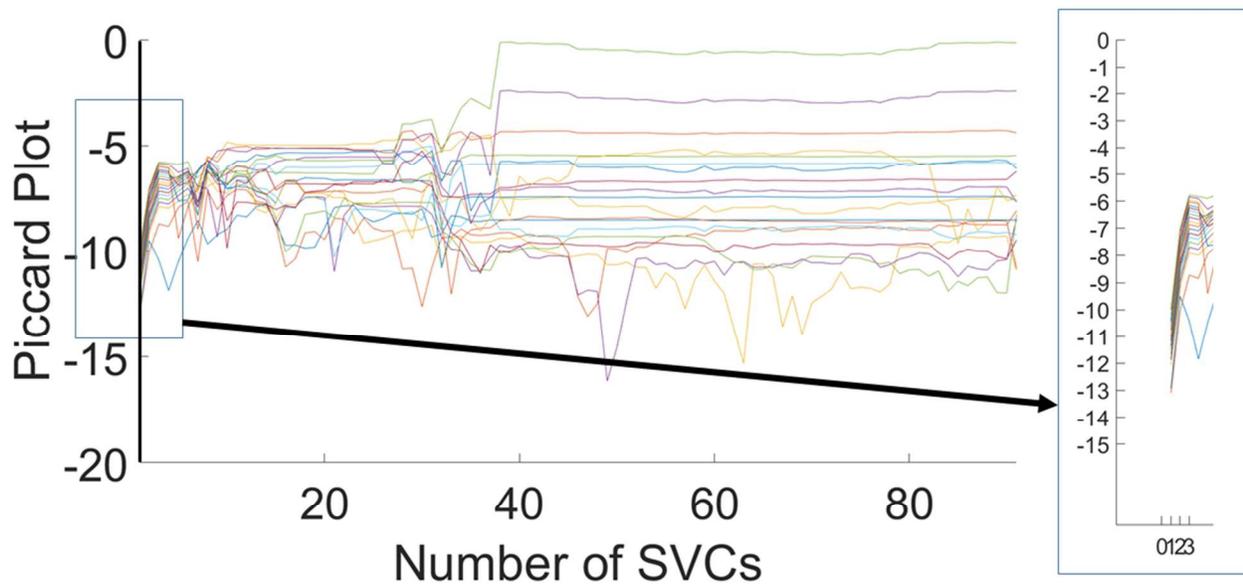


Figure S21: Experimental Data (cf. Fig. 4): Singular value cut-off on the Piccard plot for segment 1 (1 to 2.7 nm) of r region.

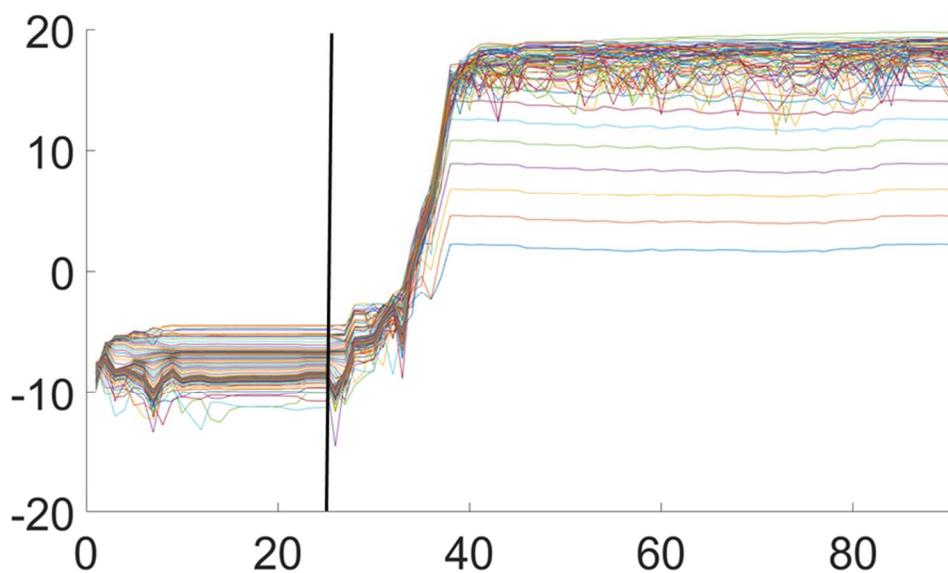


Figure S22: Experimental Data (cf. Fig. 4): Singular value cut-off on the Piccard plot for segment 2 (2.8 to 10 nm) of r region.