Extraction of Weak Spectroscopic Signals with High Fidelity: Examples from ESR

Madhur Srivastava,[*,†,‡] Boris Dzikovski,[¶,‖] and Jack H. Freed[*,¶,§]

†*Department of Chemistry and Chemical Biology, Cornell University, Ithaca, USA - 14853*

‡*National Biomedical Center for Advanced Electron Spin Resonance Technology, Cornell University, Ithaca, USA - 14853*

¶*Department of Chemistry and Chemical Biology, Cornell University, Ithaca, USA*

§*National Biomedical Center for Advanced Electron Spin Resonance Technology, Cornell University, Ithaca, USA*

‖*Current address: Bruker BioSpin, Boston, USA - 01821*

E-mail: ms2736@cornell.edu; jhf3@cornell.edu

Supporting Information

Below are contents of the supporting information:

1. NERD Algorithm

2. Discrete Wavelet Transform: Background and Mathematical Formulation

3. Wavelet Denoising: Process and Formulae, including figures

Algorithm

The algorithm (cf. Algorithm 1) for NERD is given below, and illustrated in the block diagram given in Fig. S7. The NERD software can be accessed through `denoising.cornell.edu`

Algorithm 1 NERD Algorithm

1. Select a wavelet.
2. Apply UDWT.
3. Select k ($1 \leq k \leq N$) decomposition levels to denoise the Detail components, where $N = \lfloor \log_2(Signal_{Length}) \rfloor$.
4. Select the k Detail components and the k^{th} Approximation component for noise thresholding.
5. Calculate noise thresholds each of the k Detail components and the k^{th} Approximation component.
6. Apply hard thresholding to the k selected Detail components and the k^{th} Approximation component.
7. Select signal location window using the k^{th} Detail component. In some cases one can also use the k^{th} Approximation component.
8. Apply signal location windowing to the 1 to k^{th} Detail component, except for those Detail component that contain all noise.
9. Apply signal location windowing to the k^{th} Approximation component. When Approximation component has signal region different from the k^{th} Detail component, use subjective method to independently obtain and apply signal location window.
10. Take the inverse undecimated discrete wavelet transform (IUDWT) of the resultant k Detail components and the k^{th} Approximation component.

Discrete Wavelet Transform

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,[1]

$$D_j[n] = \sum_{m=0}^{p-1} f[t_m] 2^{\frac{j}{2}} \psi[2^j t_m - n] \tag{S1}$$

, and

$$A_j[n] = \sum_{m=0}^{p-1} f[t_m] 2^{\frac{j}{2}} \phi[2^j t_m - n] \tag{S2}$$

where $f[t_m]$ is the discrete input signal, 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 the 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. That is:

$$\sum_{t_m=0}^{L-1} \psi_{j,k}[t_m] \psi_{\tilde{j},\tilde{k}}[t_m] = \begin{cases} 1, & \text{if} j = \tilde{j} \& k = \tilde{k} \\ 0, & \text{otherwise} \end{cases} \tag{S3}$$

$$\sum_{t_m=0}^{L-1} \phi_{j,k}[t_m] \psi_{j,k}[t_m] = 0 \tag{S4}$$

where L is the finite length of the scaling and wavelet function and their values are selected depending on the choice of wavelet.

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

$$f[t_m] = \sum_{k=0}^{p-1} A_{j_0}[k] \phi_{j_0,k}[t_m] + \sum_{j=1}^{j_0} \sum_{k=0}^{p-1} D_j[k] \psi_{j,k}[t_m] \tag{S5}$$

where j_0 is the maximum decomposition level from which input signal needs to be reconstructed.

Wavelet Denoising

The following are the steps used in Noise Elimination and Reduction via Denoising (NERD) to accomplish denoising.

a) **Undecimated Discrete Wavelet Transform:-** In NERD, we use the Undecimated DWT

(UDWT)[2] to achieve the maximum signal and noise resolution in the wavelet transform (cf. Fig. S1). This means that each Detail and Approximation component has the same length as that of the input signal. For instance, a signal with data length of 128 will have 7 Detail and Approximation components (from 7 decomposition levels), each having a length of 128 data points. The maximum number of decomposition levels, N, is defined as $N = log_2 p$ (p being the input signal length).[3] The UDWT improves the resolution in the wavelet domain by the preserving the input data length at all decomposition levels.

Furthermore, UDWT preserves the redundancy in the wavelet domain by keeping all the wavelet coefficients. This provides protection against the noise thresholding procedure. Even if a signal coefficient is removed by noise thresholding, the information can be easily recovered from other (redundant) coefficients, which is not possible in the DWT because each wavelet coefficient represents unique information.

For fast computation, we employ a fast UDWT that uses low and high pass filters shown in Fig. S2 to replace scaling and wavelet functions, respectively, in order to obtain the Approximation and Detail components as follows:

$$A_{j+1}[n] = \sum_{i=0}^{L-1} l[i] A_j[i+n] \tag{S6}$$

and

$$D_{j+1}[n] = \sum_{i=0}^{L-1} h[i] A_j[i+n] \tag{S7}$$

where $l[i]$ and $h[i]$ 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 functions.

As can be seen from equations S6 and S7 and illustrated in Fig. S1, the low and high pass filters are recursively applied on the Approximation component to obtain the Approximation and Detail components, respectively, at the subsequent level. For decomposition level 1, the

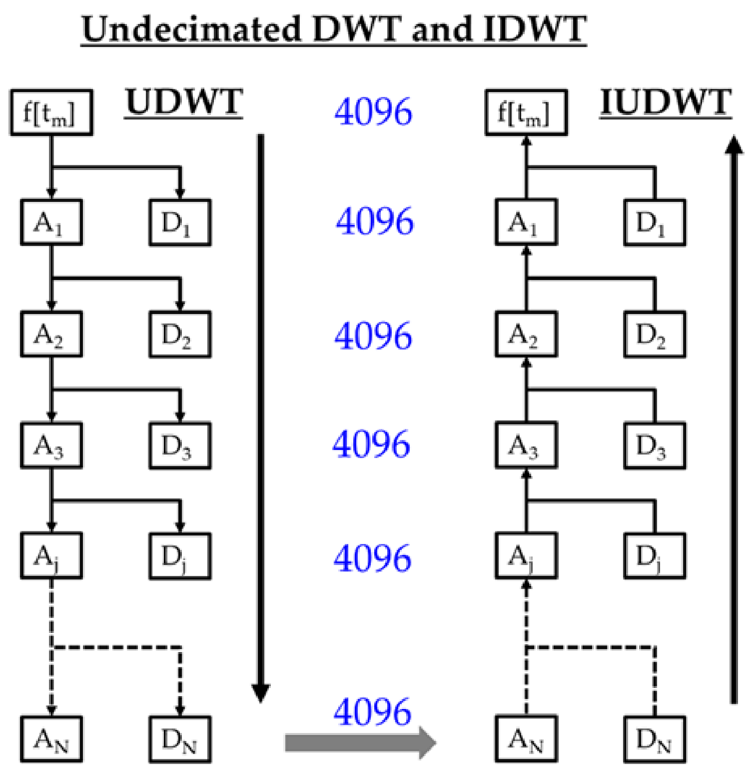

Figure S1: Block diagram of the undecimated discrete wavelet transform (UDWT) and inverse undecimated discrete wavelet transform (IUDWT). $f[t_m]$ is the discrete input signal, and D_j and A_j are the Detail and Approximation components, respectively, at decomposition level j.

input signal is the Approximation component at the 0^{th} level.

The Inverse Undecimated Wavelet Transform (IUDWT) is defined as:

$$A_j[n] = \sum_{i=0}^{L-1} \tilde{l}[i]A_{j+1}[i+n] + \sum_{i=0}^{L-1} \tilde{h}[i]D_{j+1}[i+n] \tag{S8}$$

We show in Fig. S2 the Approximation and Detail components for an ESR signal in the UDWT mode.

b) **Wavelet Selection:-** There are many standard wavelet families that can be used for the DWT. We use the coiflet family wavelet "coiflet-3" (cf. Fig. S3). 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

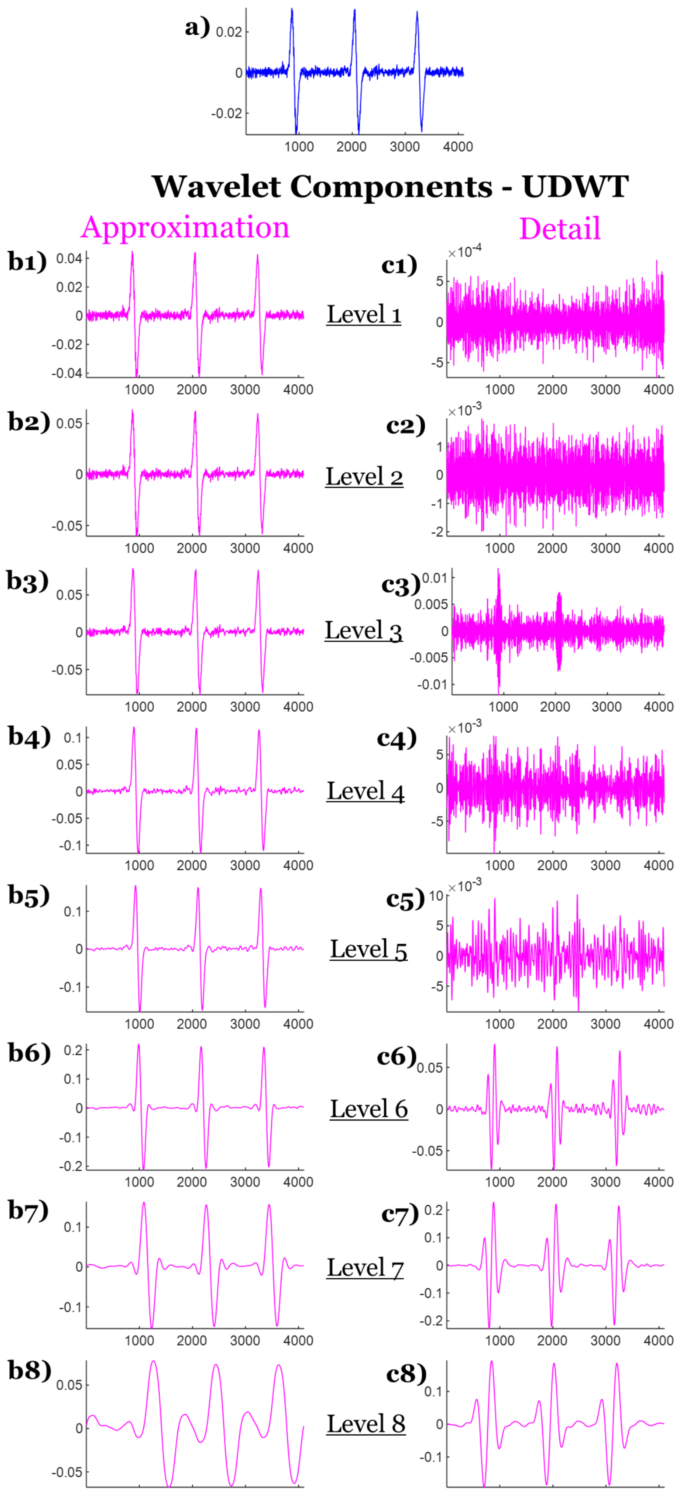

Figure S2: Approximation and Detail components from UDWT at different decomposition levels of a cw-ESR spectrum. a) Input signal is shown in subfigure with 4096 data points. Level represents the Decomposition Level.

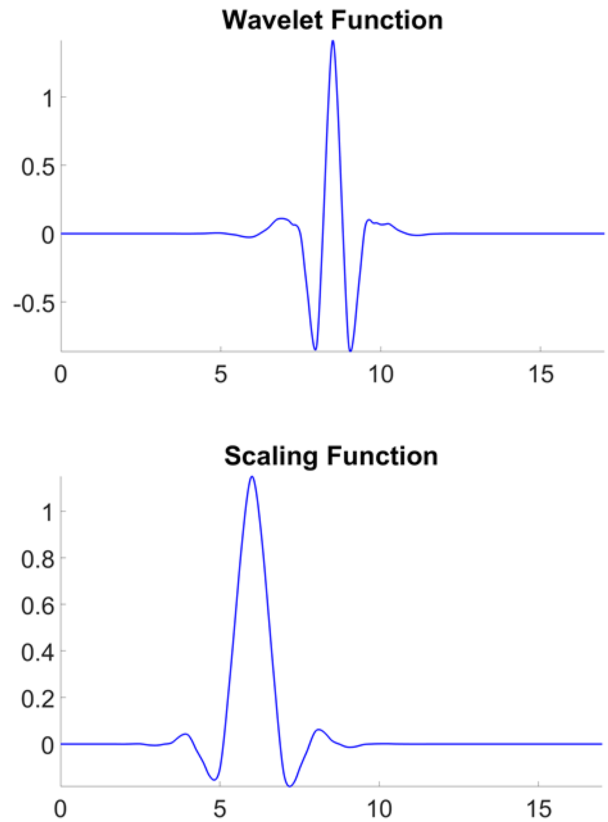

Figure S3: Wavelet and Scaling functions of Coiflet-3 wavelet.

not capture all the necessary information, whereas a larger length would yield redundant information. In Fig. S4, we show the low and high pass filter coefficients used in UDWT and its inverse for Coiflet-3.

c) **Decomposition Level Selection:-** For successful noise thresholding in the wavelet domain, all the Detail components and the highest decomposition Approximation component that contain noise need 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 determine if a Detail component contains noise, we first calculate the "sparsity" of each Detail component in the following way:[4]

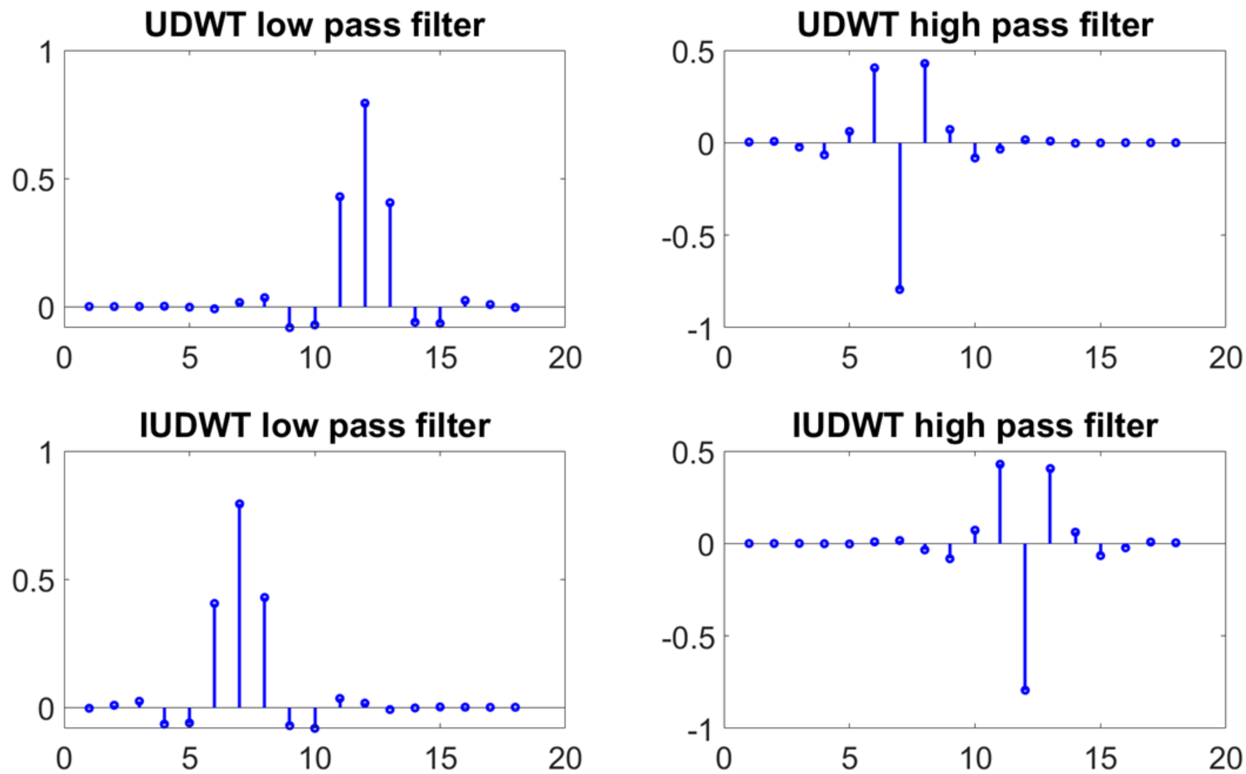

Figure S4: Low and High pass filter coefficients used in UDWT and inverse UDWT for Coiflet-3 wavelet.

$$S_j = \frac{max(|D_j|)}{\sum_{n=1}^{p} |D_j[n]|} \tag{S9}$$

where S_j is the sparsity and p is the length of the Detail component (same as the input signal for UDWT).[4] One can see 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 where $S_j < T_r$ and $S_{j+1} > T_r$. The value of T_r ranges between 0 and 1, and we have found that $T_r \approx 0.2$ is an effective cut-off; but this can be varied based on specific needs.

d) **Noise Thresholds:-** 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 (e.g. generated from instruments) have many small magnitudes, whereas

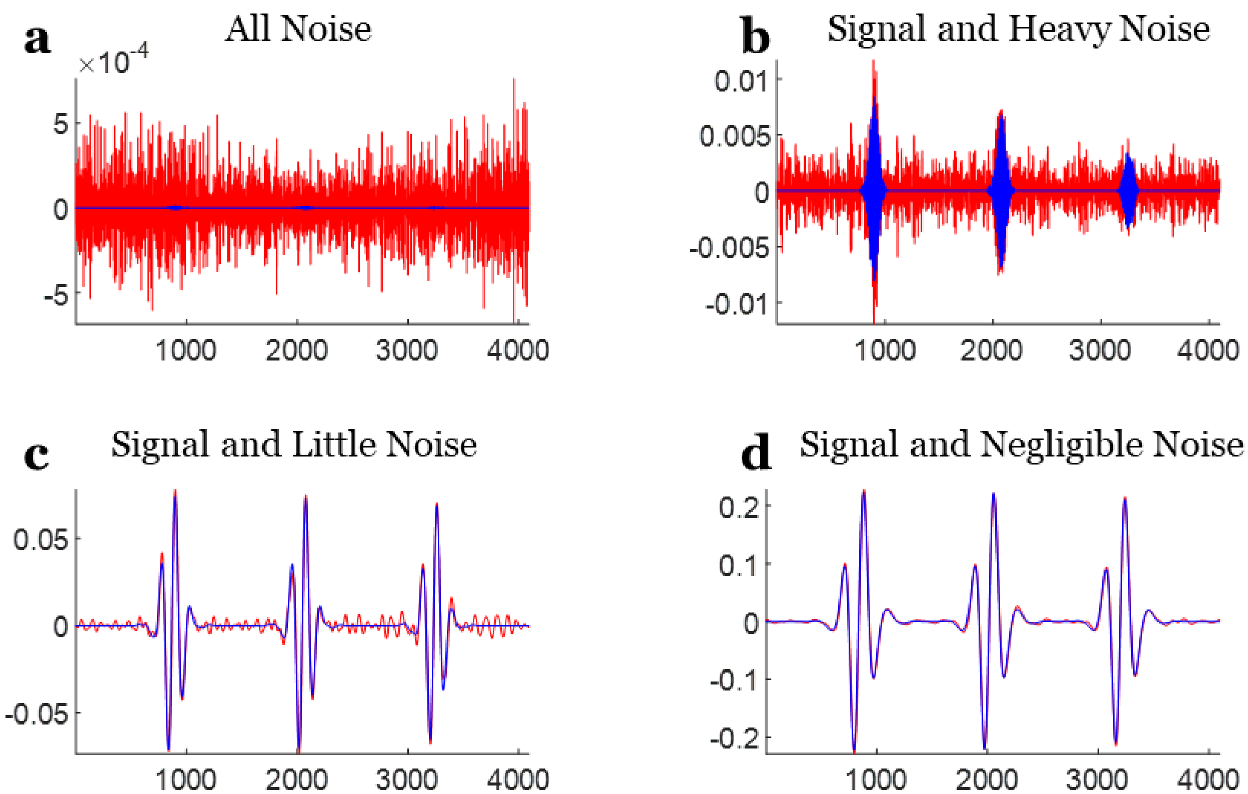

Figure S5: The four cases of the Detail component for the noisy signal: a) Detail component contains only noise, b) Detail component simultaneously contains signal and heavy noise, c) Detail component contains signal and little noise, and d) Detail component contains primarily signal with negligible noise. Input signal is shown in Fig. S2a with 4096 data points.

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 zero because they represent non-zero frequencies. The noise thresholds are calculated using the following formulae:[4]

$$\lambda_{j,L} = \mu_j - \kappa_{j,L}\sigma_j \tag{S10}$$

and

$$\lambda_{j,H} = \mu_j + \kappa_{j,H}\sigma_j \quad \text{(S11)}$$

where $\lambda_{j,L}$ and $\lambda_{j,H}$ are the lower and upper thresholds, μ_j and σ_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. The $\kappa_{j,L}$ and $\kappa_{j,H}$ values can be obtained from the formulae presented in Srivastava et al.[4]

e) **Noise Thresholding Function:-** The noise thresholds are applied in the Detail component using the hard thresholding function as follows:[4]

$$D'_j[n] = \begin{cases} 0 & : \lambda_{j,L} \leq D_j[n] \leq \lambda_{j,H} \\ D_j[n] & : otherwise \end{cases} \quad \text{(S12)}$$

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

f) **Denoising Approximation Component:-** Unlike standard denoising methods, in NERD noise thresholding is also applied to the Approximation component (as previously done in Srivastava et al.[4]). The noise thresholds are applied to the k^{th} Approximation component and are calculated in a similar manner to the Detail components using equations S10 and S11. However, the Approximation component need not be centered around 0, as it contains a d.c. component (i.e., value at 0^{th} frequency). Therefore, its lower and upper thresholds can have any value and one does not need to represent negative and positive wavelet coefficient values, respectively. The hard noise thresholding in equation S12 is modified for the Approximation component in the following way:

$$A'_k[n] = \begin{cases} a & : \lambda_{k,L} \leq A_k[n] \leq \lambda_{k,H} \\ A_k[n] & : otherwise \end{cases} \tag{S13}$$

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 are replaced. For a signal with baseline centered around 0, $a = 0$. For a non-zero baseline signal, a is assigned the baseline value.

g) **Signal Location Windowing:-** Standard wavelet denoising methods, including our previous method,[4] use noise thresholds based on the magnitude of the wavelet coefficients for each wavelet component. This approach is sufficient when the magnitude of the signal coefficients is greater than the noise coefficients as expressed below:

$$max(|w_j^{Noise} \geq 0|) < min(|w_j^{Signal} \geq 0|) \tag{S14}$$

$$max(|w_j^{Noise} < 0|) < min(|w_j^{Signal} < 0|) \tag{S15}$$

where w_j^{Noise} and w_j^{Signal} are noise and signal coefficients, respectively, of Detail or Approximation components. Equation S14 states that the magnitude of all the positive signal coefficients is larger than all the positive noise coefficients, whereas equation S15 states that the magnitude of all the negative signal coefficients is larger than all the negative noise coefficients.

The above condition also states that each wavelet coefficient is either a signal coefficient or a noise coefficient. This can be seen in Figs. S5a and S5d, where each wavelet coefficient is either signal or noise. It is particularly true for higher SNRs where the noise contribution is relatively small and is separated out in the first few Detail components. However there may be situations where both signal and noise exist in a wavelet coefficient, so it can be taken as a signal or a noise coefficient based on which one dominates, because the other's contribution to the coefficient is small. Figs. S5b and S5c illustrate this point. One can see that the few signal coefficients have very limited noise present in them, so the coefficients are overwhelmingly dominated by the signal information. Moreover, the overall contribution of a particular wavelet coefficient to the original signal is limited. Thus, its limited noise will have negligible effect on the final result.

These equations and noise thresholds based on the coefficient magnitudes also imply that one has to select a wavelet that can satisfy the above condition as well as have the signal SNR sufficiently high to ensure that wavelet coefficients representing signal are larger in magnitude than those of noise coefficients.

However, for smaller initial SNRs (e.g. of the order of unity) equations S14 and S15 will no longer be generally applicable; in such situations, signal coefficients can have magnitudes smaller than the noise coefficients. For this purpose, we introduce the concept of signal location windowing based on the coefficient location. Although the noise of the components may be larger in magnitude than signal coefficients, signal location windowing can still be utilized to recover signal coefficients using the location information. Because the signal will occur at the same location in each wavelet component, we can use cross-component information to localize where signal coefficients are occurring. We can then apply this to avoid noise thresholding those wavelet component locations and hence restore the coefficients whose signal magnitudes are smaller than the noise magnitudes. This procedure can be applied to both Detail and Approximation components. This is conveniently accomplished using the UDWT for taking wavelet transforms, since there is a one-to-one location correlation amongst the components for localizing the signal coefficients.

Subjective Approach:- For subjective analysis, we first apply noise thresholding to all the Detail components (cf. Fig. 1a) up to the maximum decomposition level k selected to denoise (cf. Fig. 1b). This removes noise coefficients and the signal coefficients whose magnitude is less than the maximum magnitude of noise coefficients. We then plot all the k

noise-thresholded Detail components (shown in Fig. 1b). The noise coefficients dominating the signal coefficients will primarily be in the initial decomposition levels as this is where noise is heavily present and signal is weak, but as we move towards the maximum decomposition level k, the signal coefficients will become dominant (cf. Fig. 1b6); therefore, the noise thresholding at that decomposition level will only remove the noise coefficients. We select the location of signal coefficients from the maximum decomposition level k, to apply signal location windowing. Once determined, we then apply the signal location window to all the Detail components where both signal and noise are present to recover the low magnitude signal coefficients (cf. Fig. 1c). Note that the results in Fig. 1c compare favorably with those from the reference, shown in Fig. 1d. The Detail components containing only noise coefficients are exempted in the signal location windowing procedure, since they do not contain signal coefficients in the first place. The noise-only Detail components are determined using the sparsity parameter S_j in equation S9 and described in Srivastava et al.[4] in detail. For practical purposes, S_j values less than 0.01 indicate noise-only.

Figs. 1a,b,c,d are in the *main* text

Objective Approach:-In the objective approach, similar to the subjective one, the Detail components are first noise thresholded up to the maximum decomposition level k. Then, the boundaries of the non-zero component values at this k^{th} decomposition level are assigned as defining the signal location window. Then windowing is applied to the Detail components that contain both signal and noise coefficients, leaving the noise-only coefficients unaffected. That is, such wavelet coefficients lying within the signal location window are restored. The all-noise Detail components are identified using the sparsity parameter S_j, as in the subjective approach. The procedure can be formulated as:

$$V_{loc} = \{n : \forall D_k[n] \neq 0\} \tag{S16}$$

$$D'_j[n] = \begin{cases} D_j[n] & : if\ n \in V_{loc} \\ D'_j[n] & : otherwise \end{cases} \tag{S17}$$

where V_{loc} is the index of all the non-zero Detail coefficients. Equation 16 yields all the locations of signal coefficients (i.e., non-zero Detail coefficients at decomposition level k). Equation 17 assigns the wavelet coefficients their original value within the vertical thresholds for decomposition levels 1 to $k-1$, except for those that are all noise Detail components.

The all-noise Detail components are determined using equation 9, where for all the noise Detail components, $S_j \leq 0.01$. (This value can be adjusted depending upon the type/class of signal and the number of data points in the input signal.)

i) **Signal Location Windowing on Approximation Component:-** The signal location windowing procedure can also be applied to the Approximation component at the k^{th} decomposition level (cf. Fig. S6). The signal location window selected for the k^{th} Detail component is utilized for restoring the signal coefficients in the Approximation component. In many cases, the noise thresholding procedure is sufficient to remove all the noise coefficients in the Approximation component (Fig. S6c), like for the k^{th} Detail component, so signal location windowing may not be necessary. Note that in those cases the signal location window can also be selected from the Approximation component and applied to the Detail components, because the signal locations for both of them are the same due to the use of UDWT. However, in the case where low frequency noise is substantial, the signal location windowing needs to be selected from the k^{th} Detail component and applied to the Approximation component to reinstate the signal coefficients. The procedure can be formulated as:

$$V_{loc} = \{n : \forall D_k[n] \neq 0\} \tag{S18}$$

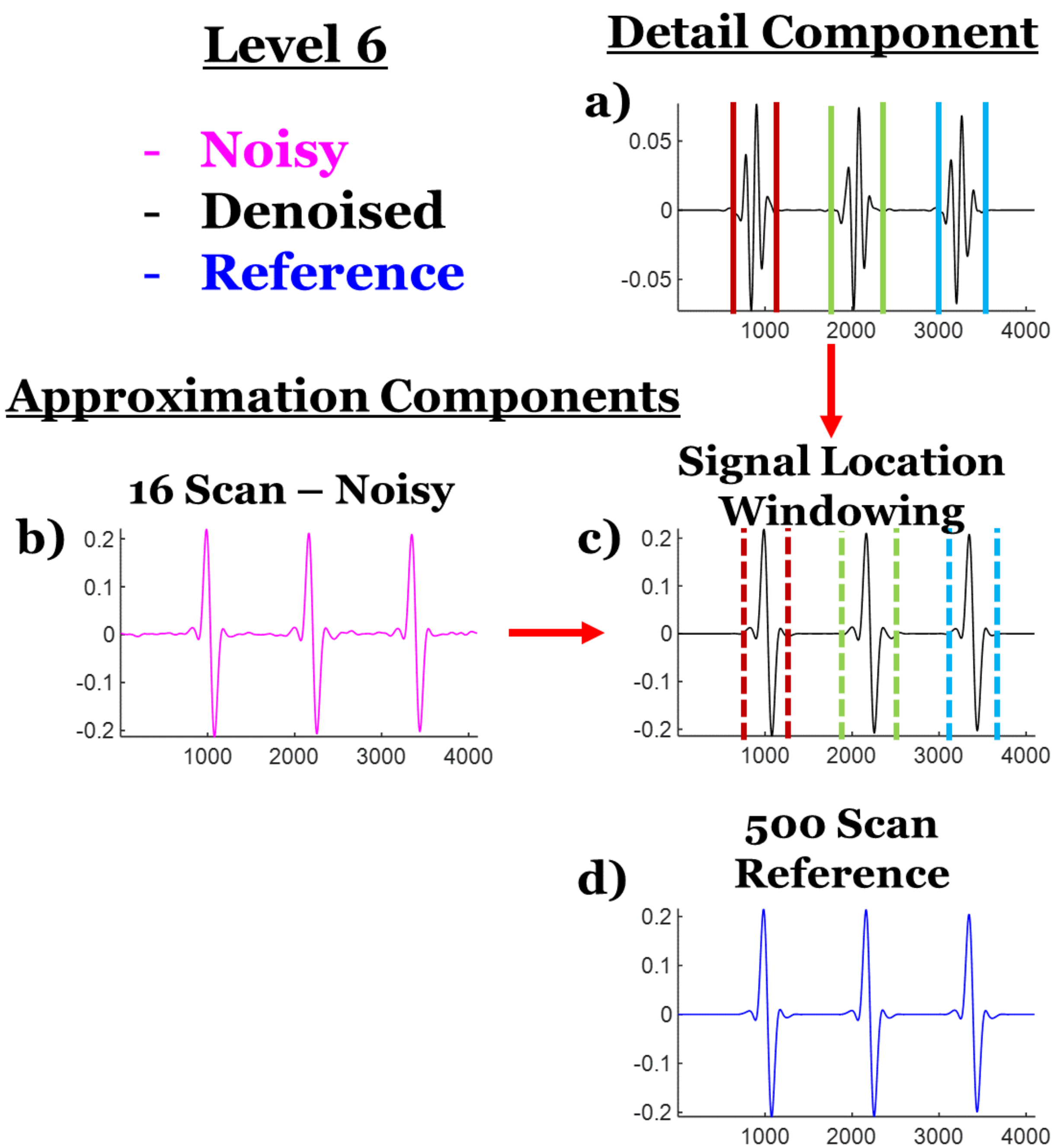

Figure S6: Example of signal location windowing in the k^{th} Approximation component (Level 6) using the window obtained from the k^{th} Detail component (Level 6).

$$A'_j[n] = \begin{cases} A_j[n] & : if\ n \in V_{loc} \\ A'_j[n] & : otherwise \end{cases} \tag{S19}$$

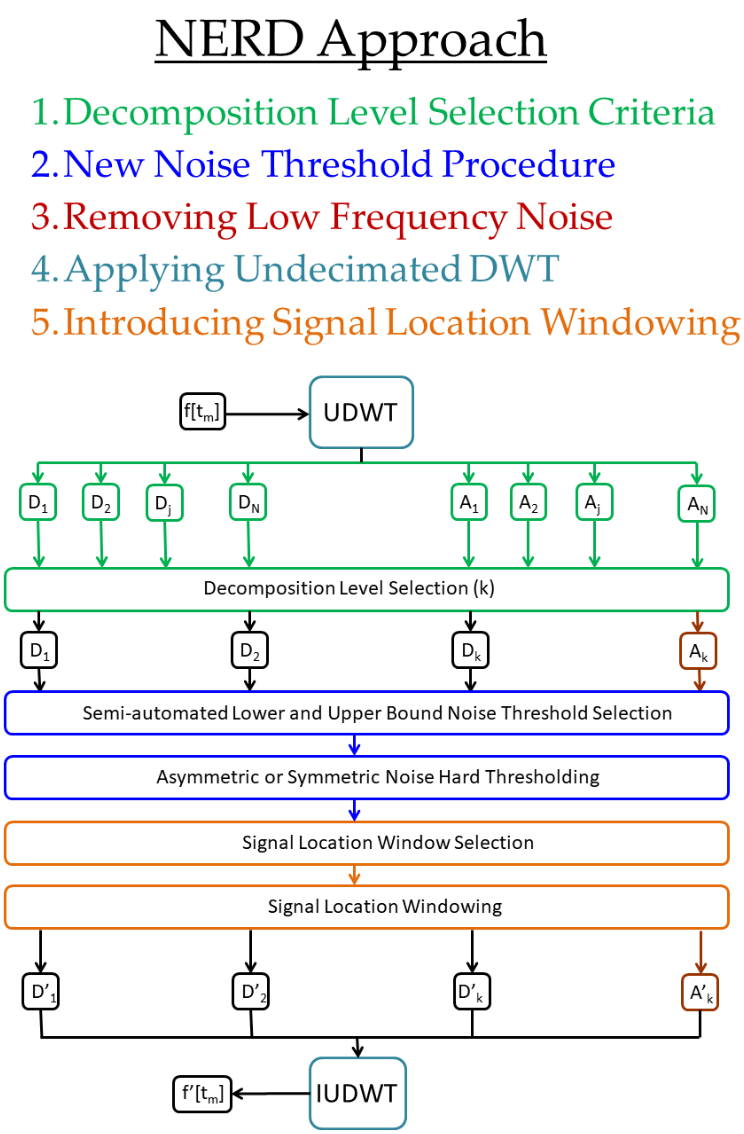

Figure S7: Block diagram of NERD Approach. $f[t_m]$ and $f'[t_m]$ are the noisy and denoised signals, respectively, and D_j and A_j are the Detail and Approximation components, respectively, at decomposition level j.

In special cases, the Approximation component may have a wider or narrower signal coefficient location window compared to the Detail component due to baseline or any other low frequency feature. In those scenarios, the subjective approach will be applied to the Approximation component.

The NERD approach is summarized in the block diagram of Fig. S7.

References

(1) Addison, P. *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance*, 2nd ed.; CRC Press: London, UK, 2016.

(2) Starck, J.; Fadili, J.; Murtagh, F. The Undecimated Wavelet Decomposition and its Reconstruction. *IEEE Trans. Image Process.* **2007**, *16*, 297–309.

(3) Antoniadis, A.; Oppenheim, G. *Wavelets and Statistics*; Springer, New York, NY, 1995.

(4) 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.