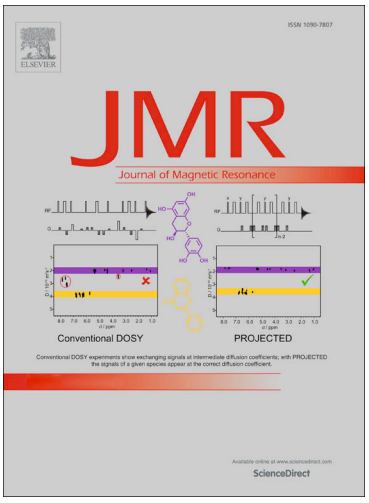
Journal Pre-proofs



DESPERATE: A Python Library for Processing and Denoising NMR Spectra

Adam R. Altenhof, Harris Mason, Robert W. Schurko

PII: S1090-7807(22)00178-1
DOI: https://doi.org/10.1016/j.jmr.2022.107320
Reference: YJMRE 107320

To appear in: *Journal of Magnetic Resonance*

Received Date: 16 August 2022
Revised Date: 4 October 2022
Accepted Date: 20 October 2022

Please cite this article as: A.R. Altenhof, H. Mason, R.W. Schurko, DESPERATE: A Python Library for Processing and Denoising NMR Spectra, *Journal of Magnetic Resonance* (2022), doi: https://doi.org/10.1016/j.jmr.2022.107320

DESPERATE: A Python Library for Processing and Denoising NMR Spectra

Adam R. Altenhof[1,2], Harris Mason[3,*], and Robert W. Schurko[1,2,*]

1. Department of Chemistry and Biochemistry, Florida State University, Tallahassee, FL 32306, USA

2. National High Magnetic Field Laboratory, 1800 East Paul Dirac Drive, Tallahassee, FL 32310, USA

3. Chemistry Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

*Author to whom correspondence should be addressed.

E-mail: hmason@lanl.gov, rschurko@fsu.edu

Tel: 505 665 2692, (850)-645-8614

Abstract

NMR spectroscopy is an inherently insensitive technique with respect to the amount of observable signal. A common element in all NMR spectra is random thermal noise that is often characterized by a signal-to-noise ratio (SNR). SNR can be generically improved experimentally with repetitive signal averaging or during post-processing with apodization; the former of which

often results in long experimental times and the latter results in the loss of spectral resolution. Denoising techniques can instead be used during post-processing to enhance SNR without compromising resolution. The most common approach relies on the singular-value decomposition (SVD) to discard noisy components of NMR data. SVD-based approaches work well, such as Cadzow and PCA, but are computationally expensive when used for large datasets that are often encountered in NMR (*e.g.,* Carr-Purcell/Meiboom-Gill and *n*D datasets). Herein, we describe the implementation of a new wavelet transform (WT) routine for the fast and robust denoising of 1D and 2D NMR spectra. Several simulated and experimental datasets are denoised with both SVD-based Cadzow or PCA and WT's, and the resulting SNR enhancements and spectral uniformity are compared. WT denoising offers similar and improved denoising compared with SVD and operates faster by several orders-of-magnitude in some cases. All denoising and processing routines used in this work are included in a free and open-source Python library called DESPERATE.

Keywords: Denoising; Signal Processing; Software; Wavelet Transform; CPMG; MQMAS

1. Introduction

Due to its low-energy nature, NMR spectroscopy is the most powerful spectroscopic technique for the identification of unique chemical environments and explorations of molecular-level dynamics; however, this comes at a steep price in terms of sensitivity (*i.e.*, the amount of observable signal from a sample). The sensitivity of NMR largely depends on the nuclear

Zeeman interaction, which is associated with energies that depend on the gyromagnetic ratio of the nucleus (γ) and the magnetic field strength (B_0), and are orders of magnitude less than thermal energy (*i.e.*, $|\hbar\omega_0| = |\hbar\gamma B_0| \ll kT$). Hence, NMR signals can be enhanced, thereby increasing the sensitivity of the NMR experiment, by working at higher fields,[1,2] lower temperatures,[3] and/or choosing to study high-γ nuclei – however, these options are not always the reasonable or desirable. High magnetic fields are extremely costly, low temperatures that substantially enhance the Boltzmann ratios to yield high signals are not accessible or practical for most spectrometers and many chemical/biological samples, and often, it is important to study the many unreceptive, low-γ, low-natural abundance nuclei from across the Periodic Table.

There are many alternate means of enhancing both signal and signal-to-noise ratios (SNR) in NMR experiments, including high-quality shims and field stabilization hardware; improved advanced designs in electronics, probes, and coils; and most recently, hyperpolarization techniques that require specialized hardware (*e.g.*, dynamic nuclear polarization, DNP).[4–6] Another option is to sensitize the NMR signal by way of radiofrequency (RF) pulse sequences, including general techniques for coherence transfer (*e.g.,* nuclear Overhauser effect,[7] cross polarization,[8–10] *etc*.), exploiting relaxation (*e.g.,* steady-state free precession,[11] Carr-Purcell/Meiboom-Gill,[12,13] paramagnetic doping,[14,15] *etc*.), indirect detection (e.g., HSQC, HMQC,[16–18] *etc*.), and many other approaches.[19–21] There is no one piece of hardware or pulse sequence that is guaranteed to provide enhancements of signals under all conditions. However, one feature universal to all NMR experiments is the presence of random uncorrelated noise, the decrease of which can greatly impact the outcome of NMR experiments and their general sensitivity.

In pulsed-Fourier transform (FT) NMR, repeated acquisitions of NMR transients allow for the summation of the coherent and correlated signals, which average against the random uncorrelated noise.[22] Subsequently, in order to achieve an N-fold increase in SNR, N^2 number of repetitions/scans are required,[23] which often leads to long experimental times. An alternative approach is to use signal processing techniques to increase the SNR. This can be accomplished using one of the simplest approaches by multiplication of time-domain FIDs with window functions or filters (*i.e.,* apodization). This has the key dilemma that increasing the SNR with apodization often comes at the cost of resolution, or vice-versa.

Advanced processing techniques can instead be used to *denoise* NMR spectra (*i.e.,* partial or complete removal of the noise). Denoising techniques often rely on statistical analysis where singular value decomposition (SVD)-based techniques have been implemented widely for denoising NMR and other spectroscopic data.[24–26] Koprivica *et al.* recently developed compressed sensing-based methods for multidimensional denoising.[27] Cadzow denoising uses time-domain data (*i.e.*, from the NMR FID) to first form a Hankel or Toeplitz matrix, and then SVD is used to factor that matrix.[28,29] A certain number of singular values and singular vectors are discarded (*i.e.*, the ones that are highly correlated with noise) and reverse factorization is performed to yield a denoised Hankel or Toeplitz matrix,[30–32] from which a denoised NMR FID can then be reconstructed and processed as normal. Principal component analysis (PCA), an SVD-based matrix factorization, has also been used for denoising 2D relaxation datasets.[33–36]

The wavelet transform (WT) is another approach for denoising that has found application in many areas of signal processing. WT's represent signals as a superposition of orthonormal basis functions called wavelets.[37,38] The traditional discrete WT (DWT) downsamples the signal to a number of decomposition levels, k, where the scaling of the wavelet is adjusted at each level.

Scaling has the benefit of analyzing highly localized frequencies in a signal, making it useful for the identification of noise and its subsequent removal.[39,40] Once the signal is represented in these downsampled and scaled decomposition levels, noise is removed with thresholding. Hard thresholding sets spectral intensities to zero that are below some threshold constant, λ, whereas soft thresholding scales spectral intensities according to some modulation of the spectral intensity, the thresholding constant, and sometimes other constants.[39–44] Sirivastava *et al*. have implemented several WT protocols that use either hard or soft thresholding for denoising electron paramagnetic resonance (EPR) spectra.[45–47] To date, WT denoising has found some usage for processing NMR data, including denoising relaxation data[48,49] and a limited number of applications to processing 2D NMR datasets.[50,51]

Herein, we explore the potential of using WTs for denoising 1D and 2D NMR data with differing lineshapes and compare their performance to SVD-based methods. Signals in NMR spectra generally appear as high-resolution peaks originating from isotropic manifestations of NMR interactions [*i.e.,* often encountered in solution-state NMR or magic-angle spinning (MAS) solid-state NMR (SSNMR)] or as patterns with large frequency dispersions originating from anisotropic interactions or other sources of inhomogeneous broadening (*i.e.,* static or slow-MAS SSNMR spectra). WT denoising is applied to both static and MAS SSNMR spectra and demonstrated to be robust for spectra with sharp peaks and anisotropically broadened powder patterns. 2D multiple-quantum MAS (MQMAS) spectra,[52,53] which feature both isotropic peaks and anisotropic dispersions, are also processed with WT denoising. The performance of WT denoising is comparable or superior to that of standard SVD-based approaches and operates with decreased computation times, in some cases, by several orders of magnitude. Careful consideration is given to the appearance and uniformity of peaks and patterns in the denoised

spectra via assessment with SNR measurements and the differences in structural similarity indices (SSIM, *vide infra*). The WT denoising routines are implemented, along with other useful processing functions, in a free and open-source Python library called DEnoising SPEctRA in pyThon with wavElets (DESPERATE).

2. Methods

2.1 Samples

Tin(II) oxide [SnO, Sigma Aldrich], $Pt(NH_3)_4Cl_2 \cdot H_2O$ [Sigma Aldrich], sodium chloride [NaCl, Sigma Aldrich], sodium sulfate [Na_2SO_4, Sigma Aldrich], and rubidium nitrate [$RbNO_3$, Sigma Aldrich] were purchased and used in all NMR experiments without further purification. The identities and purities of the samples were verified through comparisons with previously reported NMR spectra and PXRD patterns.[54–56] All samples were ground into fine powders and packed into 3.2 mm rotors.

2.2 Solid-State NMR Spectroscopy

NMR spectra were acquired using a Bruker Avance NEO console and a 14.1 T Magnex/Bruker ($\upsilon_0(^1H)$ = 600 MHz) wide-bore magnet at resonance frequencies of $\upsilon_0(^{119}Sn)$ = 223.772 MHz, $\upsilon_0(^{195}Pt)$ = 129.001 MHz, $\upsilon_0(^{23}Na)$ = 158.738 MHz, and $\upsilon_0(^{87}Rb)$ = 196.348 MHz. A home-built 3.2 mm triple resonance (HXY) magic-angle spinning (MAS) probe was used for all experiments. Spectra were acquired with ^1H continuous-wave (CW) decoupling with RF fields of 50 kHz for compounds having protons. RF pulse powers and chemical-shift reference frequencies were calibrated using the following standards: ^{119}Sn reference: $Sn(CH_3)_4$ (*l*) with δ_{iso} = 0.0 ppm; ^{195}Pt reference: 1.0 M Na_2PtCl_6 (*aq*) with δ_{iso} = 0.0 ppm; and the following were only

used as chemical-shift references: ^{23}Na reference: NaCl (*s*) with δ_{iso} = 0.0 ppm; ^{87}Rb reference: 0.1 M RbCl in D_2O (*aq*) with δ_{iso} = 0.0 ppm. ^{23}Na and ^{87}Rb RF pulse powers were calibrated by finding the main spin-lock rotary resonance conditions, $(S+1/2)\nu_1 = \nu_{rot}$,[57] at ν_{rot} = 5 kHz and 10 kHz with Na_2SO_4 and $RbNO_3$, respectively.

2.3 Computational methods

All numerical simulations were conducted in SIMPSON 4.2.1[58–60] using either 4180 or 28656 orientations sampled according to the ZCW averaging scheme.[61] CPMG experiments were simulated using ideal pulses with 50 or 300 spin echoes. MQMAS experiments were simulated using a three-pulse whole-echo style sequence[62] using ideal pulses and matrix filtration for coherence selection.

All processing and denoising routines were performed using a custom in-house Python 3 library called DESPERATE. The library can be found at github.com/rschurko/desperate and pypi.org/project/desperate. It has dependencies on NUMPY,[63] SciPy,[64] PyWavelets,[65] and Matplotlib.[66]

All simulations and processing were performed on a PC operating with Windows 10 using an Intel i9-9920X CPU.

2.4 Denoising Techniques

In this work, denoising is accomplished with either SVD-based techniques or using wavelet transforms (WT). SVD approaches include either the Cadzow denoising technique on time-domain signal or running PCA of 2D time- or frequency-domain signal.[28,29,33,34] Cadzow denoising has been described in detail elsewhere[30–32] and has been implemented in DESPERATE

by running SVD on the Hankel matrix of the time-domain NMR FID. The number of discarded singular values is determined automatically using a derivative of the singular values (**Figure S1**). PCA has also been implemented and should, in principle, be used for processing 2D NMR data.

Herein, we have implemented a WT method for denoising frequency-domain 1D and 2D NMR data that operates with the following steps (**Scheme 1**):

1. Preprocess the 1D or 2D NMR FID, including the application of window functions, echo-coaddition, zero-filling, FFT, phase correction, and/or shearing, *etc*.
2. Evaluate the stationary WT (SWT) of the real component of the complex frequency-domain data to some number k of decomposition levels, d_k, that each contain an approximation component and detail component, A_k and D_k, respectively.
3. Use signal windowing to isolate the baseline noise in every A_k.
4. Perform a modified soft-thresholding routine on *all* A_k and D_k using a threshold constant, λ, determined from the windowed noise per decomposition level.
5. Evaluate the inverse-SWT (ISWT) to reconstruct the denoised NMR spectrum.

The SWT returns undecimated values of A_k and D_k whereas a discrete WT (DWT) decimates the components from down sampling. The biorthogonal 2.2 (bior2.2) wavelet is used for the SWT by default,[65] but others can be used in the WT method above. A modified soft-thresholding routine is used as described by Wang and Dai,[67] where

$$d_i = \begin{cases} d_i - \alpha\dfrac{\lambda_i^4}{d_i^3}, & |d_i| \geq \lambda_i \\ (1-\alpha)\dfrac{d_i^5}{\lambda_i^4}, & |d_i| < \lambda_i \end{cases} \quad (1)$$

for each ith decomposition level up to $i = k$, where d_i represents both the approximation and detail components, $\lambda = \sigma(noise)\sqrt{2\log(n)}$ for a component with n number of points, and $\alpha \in$

[0,1] (α = 0 by default). For SWTs, n must be divisible by 2, *i.e.,* $n = 2^x$, where the maximum possible $k = x$, and k is manually set by the user.

Denoising in the frequency domain, as opposed to the time domain, has the benefits of easy identification of baseline noise with or without signal windowing (*i.e.,* the signal and noise are defined as binary regions using a baseline derivative and in an FFT-shifted spectrum the noise can be easily isolated or sampled since it is dispersed to the edges of the window, **Figure S2**); artificially increasing the resolution of the signal via zero-filling, which also aids in signal windowing; and the flexibility to allow the selection of more k levels by adjusting zero-filling (the maximum k cannot be changed using fixed-size time-domain data unless zero padding is used). Similar signal windowing has been used in the NERD WT denoising algorithm introduced by Srivastava *et al.*;[47] however, our approach differs by (i) the use of windowing for isolating the noise for calculating λ at each decomposition level and (ii) thresholding all values of values of A_i and D_i (from $i = 1 \ldots k$).

One clear advantage of using WTs over SVD-based approaches is computational time. For a dataset with n number of points, the corresponding Hankel matrix for Cadzow denoising in this work has dimensions $a \times b$, where $a = n/2 + 1$ and $b = n/2$; the subsequent SVD has a computational complexity of $O(ab^2)$ whereas that of the WT is $O(n)$.[25,51] An example of this is illustrated by calculating the computational times required to perform Cadzow and WT denoising routines with different numbers of decomposition levels for time- and frequency-domain data of varying sizes (**Figure S3**).

2.5 Metrics

Several metrics are used to measure and benchmark the performance of denoising techniques. Signal-to-noise ratios are defined as the maximum value of the real frequency-

domain spectrum over the standard deviation of the noise: $SNR = \max(spectrum)/\sigma(noise)$. The peak-to-peak SNR ($SNR_p$) is defined as the difference between the highest and lowest intensity 'peaks' or parts of the real frequency-domain spectrum over the standard deviation of the noise: $SNR_p = [\max(spectrum) - \min(spectrum)]/\sigma(noise)$. The structural similarity index (SSIM) is used to measure the uniformity of the spectrum or pattern as compared to some ground truth noise-free measurement either from a simulation without noise, or from an experimental spectrum with high signal averaging. SSIM is defined as[45–47,68]

$$\mathrm{SSIM}(X,Y) = \frac{(2\mu_X\mu_Y)(2\sigma_{XY})}{(\mu_X^2 + \mu_Y^2)(\sigma_X^2 + \sigma_Y^2)} - c \quad (2)$$

where Y is the reference spectrum that is considered noise-free (or has very little noise), X is the denoised spectrum, μ is the mean of the signal, σ is the standard deviation of the signal, and c is a small constant that ensures the SSIM is bound over a range of [-1, 1] and is calculated with $c = \mathrm{SSIM}(Y,Y) - 1$. A value of SSIM = 1 indicates that the two signals are identical.

3. Results and Discussion

3.1 Overview

In the following sections, denoising techniques are compared with one another using 1D and 2D SSNMR spectra (both simulated and experimental). The standard SVD/Cadzow approach is compared against our WT method for denoising. In every case, we show arrays of spectra that have different amounts of noise and corresponding *SNR* and SNR_p values as input for denoising, and comparisons of the *SNR*, SNR_p, and/or *SSIM*'s obtained from each denoising technique. WT methods are applied to (i) simulated and experimental CPMG datasets subjected to different types of processing (*i.e.*, echo coaddition and spikelet representations); (ii)

experimental wideline MAS SSNMR spectra; and (iii) experimental and simulated 2D MQMAS spectra.

3.2 1D Simulations

The performance of denoising of 1D NMR data is first benchmarked using synthetic CPMG data of a simulated solid-state static powder pattern associated with an axially symmetric chemical shielding (CS) tensor. This pattern is a good test case since the *SNR* varies across the pattern breadth, with the high- and low-frequency limits having relatively high and low *SNR*s, respectively, with intermediate *SNR* values captured in between these limits; here, the SNR_p metric is used to reflect these *SNR* differences. There are two primary ways of processing CPMG datasets that can affect the choice of denoising technique, including (i) FT of the entire CPMG echo train, revealing a "spikelet" representation of the pattern, or (ii) coaddition of the echoes into a single echo, which is subsequently Fourier transformed to yield a standard NMR pattern. The key difference is that the echo train has some n number of points, whereas the corresponding coadded echo train has n/N number of points, where N is the number of CPMG loops. We note that the latter case represents a dramatic reduction in computational time for Cadzow denoising, and as such, we do not include corresponding benchmarks for 1D CPMG spikelet spectra (*cf.* **Figure S3** and **Figure S4**). To aid in this type of processing we have included several processing routines in DESPERATE, including automatic CPMG echo coaddition, automatic phasing up-to second order, and other useful features.

Synthetic CPMG datasets used in **Figure 1** are generated and processed with the following progression. CPMG echo trains are simulated in SIMPSON by calculating 512 complex points for each spin echo over a total of 50 CPMG loops resulting in an FID with 25600

total complex points. The echo train is then multiplied by an exponential decay function with an effective T_2 constant to help mimic experimental datasets. At this point, pseudorandom Gaussian noise is added to the time-domain signal with an intensity that is set as a percentage of the maximum intensity of the signal; adjusting this percentage changes the SNR. Then, all of the echoes are coadded to give a single echo with 512 points. The coadded echo FIDs and corresponding spectra of varying SNR are regarded as the "input" data that are analyzed prior to denoising. At this point, Cadzow denoising may be used with the time-domain data. Subsequent zero-filling, FT, and phase correction then yields the NMR spectrum. If Cadzow denoising is not used, the post-FT and processed spectrum can be used as input for WT denoising instead. All coadded synthetic CPMG spectra have 16384 points. The SNR_p's and *SSIM*'s are all measured for the input, Cadzow denoised, and WT denoised spectra, and are denoted as $SNR_{p,in}$, $SNR_{p,C}$, $SNR_{p,WT}$, $SSIM_{in}$, $SSIM_C$, $SSIM_{WT}$, respectively.

Benchmarking is first presented for synthetic coadded CPMG echo trains (**Figure 1**). The input spectra resulting from the addition of different amounts of noise to the CPMG FID and subsequent echo coaddition and FT are pictured in the left column of **Figure 1** (the values of $SNR_{p,in}$ are given in **Table 1**). These same input FIDs are denoised with the Cadzow routine, multiplied by a Gaussian window function to help attenuate sinc-artifacts (the input coadded FID is also multiplied by an identical Gaussian function, **Figure S5**), and processed with FT and phase correction (**Figure 1** – middle column). $SNR_{p,C}$ increases dramatically compared to $SNR_{p,in}$ in every case (**Table 1**); however, these gains should be interpreted with caution, as significant spectral distortions are observed in most cases (*cf.* **Figure 1d-h**). It is perhaps more beneficial to compare the *SSIM*s with that of a reference spectrum that has no added noise (*i.e., SSIM* = 1.0000 indicates an identical match between the denoised spectrum and the noiseless input

spectrum). $SSIM_C$ is higher than $SSIM_{in}$ for spectra in **Figure 1a-d** (**Table 1a-d**) but drops off significantly for those in **Figure 1f-h** (**Table 1f-h**) (*i.e., SSIM* < 0.9900 and is lower than $SSIM_{in}$), which corresponds to $SNR_{p,in}$ values below *ca.* 70. This suggests a minimum $SNR_p \cong 70$ may be necessary to properly utilize Cadzow denoising in this case.

The spectra resulting from similar processing and WT denoising of the frequency-domain spectra with $k = 7$ decomposition levels (**Figure 1** – right column, **Table 1**) reveal that in every case, the $SNR_{p,WT}$ is much higher than $SNR_{p,in}$ (**Table 1**), though not to the same degree as $SNR_{p,C}$. The values of $SSIM_{WT}$ are higher than $SSIM_{in}$ values in every case except those of the spectra shown in **Figure 1g,h** (**Table 1g,h**), and also higher than all values of $SSIM_C$ except that of the spectrum in **Figure 1h** (**Table 1h**). A qualitative examination of the spectra also reveal that patterns can be partially recovered when the input SNR_p values are ≅ 63 or 57 (**Figure 1f,g**) as evidenced by the higher uniformity of the CSA patterns in comparison to those arising from Cadzow denoising. These results suggest that WT denoising may be used to recover powder patterns with $SNR_p \gtrapprox 60$ in this case. Examples of the detail and approximation components before and after thresholding are available in the supporting information (**Figure S6**).

A comparison of input and WT denoised CPMG spikelet spectra are shown in **Figure 2** (*N.B.*: in this case, a smaller CS span was used in simulations to help better visualize the spikelets, and Cadzow denoising was not benchmarked due to the lengthy time requirements, see **Figure S3** and **Figure S4**). Here, the echo train is simulated with 300 echoes that have 512 points each, multiplied by an exponential decay with an effective T_2 constant, noise is added, then the FID zero filled to 262144 points and Fourier transformed without any additional apodization, and then phase corrected. WT denoising with $k = 3$ results in spectra with increased $SNR_{p,WT}$ in comparison to $SNR_{p,in}$. The $SSIM_{WT}$'s are higher than $SSIM_{in}$ for every spectrum

except the first two (**Figure 2a,b** and **Table 2a,b**). The discrepancy between the first two spectra and the rest is the result of WT denoising removing inherent sinc artifacts and small baseline offsets in CPMG data sets (**Figure S7**). By comparison, the *SSIM* values measured from the low-frequency halves of the spectra that span from *ca.* –30 to – 90 kHz ($SSIM_{low}$ – **Table 2**), where there are less intrinsic baseline offsets and sinc artifacts, reveals dramatic increases in $SSIM_{WT,low}$ in comparison to $SSIM_{in,low}$ in most cases. A closer examination of the denoised CPMG spikelets (**Figure 2** - inset) reveals recovery of spikelets with Lorentzian-like shapes and very minor artifacts from residual noise. It is important to note that a relatively smaller number of decomposition levels, $k = 3$, and a large zero fill (*i.e.*, 262144) is optimal for WT denoising of spikelet data in comparison to that of the coadded spectra, which requires $k = 7$. This may be due to the sparse representation of the data with spikelets. The optimal choice of k can be determined empirically by comparing the resulting *SNR's*, *SSIM's*, and qualitative appearances of the spectra (*e.g.,* pattern uniformity and minimal artifacts); at a minimum, this is the only free variable that needs to be adjusted with WT denoising. The fast computational times associated with WT denoising makes the adjustment of k as simple as adjustment of line-broadening coefficients in standard apodization processing routines.

3.3 1D Experiments

^{119}Sn NMR spectra of SnO were acquired for the purpose of providing an experimental comparison to the aforementioned simulations. 1D WURST-CPMG[69] spectra were acquired with varying number of scans to systematically change the SNR_p of the data. The data was first processed with echo coaddition, resulting in an FID size of 290 points for Cadzow denoising, and spectra with 8192 points after zero-filling (**Figure 3** and **Table 3**). The input spectrum

acquired with 8 scans (**Figure 3h** – left column) has an $SSIM_{in} = 0.9878$ as compared to the spectrum acquired with 1024 scans (**Figure 3a** - left column); the former may represent a limit at which the spectrum is difficult to analyze and characterize (*i.e.,* $SSIM_{in} \lesssim 0.9900$). Cadzow denoising again offers a substantial gain in SNR_p as well as improved $SSIM_C$ in every case (**Table 3**). The Cadzow denoised spectra all feature sinc artifacts at the edges of the patterns, which result from the amplification of signal at the edges of the echo after Cadzow denoising (**Figure S5**). Again, these artifacts can be attenuated by multiplying the coadded echo with an appropriately positioned Gaussian (or some other) window function; however, this can cause a loss of resolution if the standard deviation of the Gaussian is too high. WT denoising with $k = 7$ increases $SNR_{p,WT}$ over $SNR_{p,in}$ and improves the $SSIM_{WT}$ over $SSIM_{in}$ and $SSIM_C$ (except for **Figure 3g**), and does so without introducing any sinc artifacts. Cadzow and WT denoising both result in values of $SSIM > 0.9900$ for the 8-scan dataset (**Figure 3h**), but a higher value of $SSIM_{WT} = 0.9974$ is achieved with WT denoising. The input, Cadzow denoised, and WT denoised spectra are all processed with the same Gaussian function for consistency; however, these results suggest that WT can denoise data while also using less apodization, thereby providing superior resolution to Cadzow denoised spectra.

The same experimental dataset was processed without echo coaddition, resulting in CPMG spikelet spectra (**Figure 4** and **Table 4**). The experimental CPMG echo train was zero-filled to 32768 points and $k = 3$ decomposition levels were used. The $SNR_{p,WT}$ gains are substantial, but are not as high as in analogous simulated data (*cf.* **Table 2**). The $SSIM_{WT}$ values are again lower than the $SSIM_{in}$ values, due to removal of intrinsic sinc artifacts in the CPMG spikelet spectra by WT denoising as demonstrated with $SSIM_{low}$ measurements for synthetic data (*c.f.* **Figure 2**, **Table 2**, and **Figure S7**); this is a great benefit for ensuring a uniform pattern

shape, but lowers the $SSIM_{WT}$. A closer look at the spikelets qualitatively illustrates the high-fidelity recovery of the Lorentzian-like spikelets (**Figure 4** – inset), but some line broadening is noticeable. The apparent line broadening is still far less severe than that resulting from removal of the same amount of noise with apodization. The efficient denoising of spikelet data is a great test case, as it is an analogue to denoising a high-resolution spectrum consisting of many isotropic, Lorentzian peaks, such as those observed in solution-state NMR or fast-MAS SSNMR spectra, or even in spinning-sideband (SSB) manifolds (*vide infra*).

Another 1D test case is the ^{195}Pt WURST-CPMG/MAS NMR of $Pt(NH_3)_4Cl_2 \cdot H_2O$. The SSB manifold of this pattern spans nearly 1 MHz under slow MAS (ν_{rot} = 10 kHz) at 14.1 T (**Figure 5** and **Table 5**).[70] The CPMG echo train is coadded in this case to properly reveal the SSB manifold (here, spikelet representations are not useful).[71] The coadded FID's have 3606 points for use in Cadzow denoising and the zero-filled spectra have 16384 points for WT denoising, where $k = 2$ is used. This is another good test case, since we can investigate the effects of different denoising routines on patterns that are anisotropically broadened but also possess high-resolution features. Substantial SNR_p gains are afforded from both Cadzow and WT denoising (**Table 5**). There is no clear superior performance indicated by comparison of $SSIM_C$ and $SSIM_{WT}$ across all of the spectra, as compared against the input spectrum acquired with the highest signal averaging (2048 scans, **Figure 5a**). Both techniques appear to fail to recover reliable spectra when the $SNR_{p,in} \lesssim 55$ (**Figure 5e,f**), which agrees with benchmarks established by simulations. We note that since many data points need to be recorded for high-resolution MAS data, that WT denoising may offer better scalability in terms of reducing computation times.

A final 1D test case is presented to simultaneously test denoising of isotropic and broad anisotropic spectral features. The ^{23}Na NMR of a NaCl:Na_2SO_4 3:50 w/w mixture was acquired with Bloch decay at ν_{rot} = 10 kHz and features a single sharp peak for NaCl and a comparatively broader anisotropic powder pattern for Na_2SO_4 (**Figure 6**) due to the relatively small and large quadrupolar couplings, C_Q, respectively. The rotor was half-filled with the mixture and half-filled with adamantane to decrease the sample's receptivity. The FIDs have 1978 points for use in Cadzow denoising and the zero-filled spectra have 8192 points for WT denoising, where k = 5 is used. Previous 1D tests (*vide supra*) reveal an optimal k = 7 or 2-3 for broad anisotropic-type spectra or sparse isotropic-type spectra (*i.e.,* series of spikelets or SSBs), respectively. Therefore, an intermediate value of k = 5 used herein may be optimal when both types of features arise in the same spectrum. Cadzow and WT denoising perform comparatively well except for the lowest SNR experiments (**Figure 6g,h** and **Table 6**). For the spectra acquired in 2 scans with $SNR_{p,in}$ = 46 (**Figure 6g**) the denoised spectra have comparable SNR_p's and *SSIM*'s (**Table 6**); however, the Cadzow denoised spectrum shows an artifact at *ca.* – 12 ppm and reduced resolution in the Na_2SO_4 pattern at *ca.* – 26 ppm, whereas the WT denoised spectrum maintains the pattern resolution. The denoised spectrum acquired with 1 scan has $SNR_{p,in}$ = 33 (**Figure 6h**) where WT denoising appears to maintain the pattern resolution better than Cadzow; however, the WT denoised spectrum reveals a distorted baseline as reflected by $SSIM_{WT}$ = 0.9978 being lower than $SSIM_C$ = 0.9984. These results approximately agree with other 1D data, where a lower limit for using WT denoising is around $SNR_{p,in} \lesssim 46$ in this case.

3.4 2D Simulations

An exciting aspect of using WT for denoising is its fast and easy implementation for *n*D NMR data. Herein, we test the 2D WT denoising routine on a synthetic MQMAS dataset with varying amounts of noise. MQMAS datasets represent ideal test cases as denoising can be simultaneously tested for direct (F_2) and indirect (F_1) dimensions, which respectively represent anisotropic and isotropic patterns. F_1 is properly scaled for the isotropic chemical shift axis (it is sometimes referred to as F_{iso}, but herein, is called F_1 for simplicity).[72] For 2D MQMAS data, the *SNR* (*N.B.:* not the SNR_p) is reported for F_2 and F_1 dimensions for simplicity (referred to as SNR^{F2} and SNR^{F1}, respectively), along with the *SSIM*s for the anisotropic slices extracted along the F_1 axis (referred to as $SSIM^i$, i = 1, 2, and 3 for the F_1 spectra from top to bottom in the order they are displayed). Herein, ^{87}Rb MQMAS datasets for $RbNO_3$ are simulated in the time domain, with B_0 = 14.1 T and ν_{rot} = 10 kHz, using reported EFG and CS tensor parameters for the three ^{87}Rb sites.[56] The time-domain data are simulated with 1024 × 128 points ($t_2 \times t_1$) and then noise is added with varying intensity to cause corresponding varying *SNR*'s in the different input spectra, and then are zero-filled to 4096 × 512 points ($F_2 \times F_1$) prior to generating the frequency-domain spectra. All MQMAS spectra are displayed with the F_2 projection spectrum as a skyline of the maximum intensity over all of the F_1 slices, while the F_1 projection spectrum is shown as a sum of all F_2 slices, to properly capture the relative F_1 peak intensities (**Figure S8**). N.B.: a forthcoming study will highlight the simultaneous usage of 2D WT denoising and non-uniform sampling for various multidimensional NMR experiments.

The first example input spectrum has a moderate SNR_{in}^{F1} and SNR_{in}^{F2} (**Figure 7a** and **Table 7**). 2D Cadzow denoising was tested by running the routine over each time-domain t_1 slice individually. The resulting spectrum shows no increase in SNR_C^{F1} and a small increase in SNR_C^{F2}, with some artifacts/defects noticeable in the F_1 slices (**Figure 7b** and **Table 7**). Cadzow

denoising is time consuming even when applied to a single FID; therefore, incremented Cadzow over a number of t_1 slices is very computationally expensive. A better alternative is to apply SVD directly on the 2D spectral matrix; here, this is implemented with PCA denoising where three principal components are retained for denoising (**Figure 7c**). PCA does a good job of increasing SNR_{PCA}^{F1} over SNR_{in}^{F1}; however, SNR_{PCA}^{F2} decreases, and no significant differences are observed in the F_1 slices between the input and PCA denoised spectra. This agrees with previous results for PCA denoising of relaxation data, where the SNR is enhanced in the indirect 'relaxation dimension' rather than in the frequency dimension (F_2).[33]

2D WT denoising is implemented in a similar manner as described earlier (**Scheme 1**), where the 2D SWT operation instead yields four components in each decomposition level: one approximation and three detail; in addition, a 2D signal windowing algorithm is used (**Figure S9**). WT denoising uses $k = 2$ decomposition levels for all MQMAS datasets. For the simulated MQMAS data, 2D WT increases both SNR_{WT}^{F1} and SNR_{WT}^{F2}, and increases the $SSIM_{WT}$ of the F_1 slices, where the *SSIM* is now compared to the noiseless F_1 spectra as references (**Figure 7d**, **Table 7**, and **Table 8** - row 7d). Further benchmarks are only carried out with WT denoising, since it outperforms the other 2D denoising techniques in terms of *SNR* enhancements over both axes and computational costs.

MQMAS datasets were simulated with various amounts of noise and were denoised with 2D WT's in every case (**Table 8**). Two examples are shown with relatively low SNR^{F1} and SNR^{F2} (**Figure 8**). As SNR_{in} decreases over both dimensions (*c.f.* **Figure 7**), WT yields enhanced SNR_{WT}, but also enhanced $SSIM_{WT}$'s for all but the second F_{iso}-slice spectrum (**Figure 8a,b** and **Table 8**). For even lower SNR_{in}'s (**Figure 8c,d**), the $SSIM_{WT}$s are not enhanced above *ca.* 0.9900, which indicates that the patterns may not be fully recovered (as benchmarked from 1D

examples). Visual examination of the corresponding F_1-slice spectra shows that the lower *SNR* discontinuities at the edges of the pattern are not completely recovered after WT denoising. These results suggest that when $SNR^{F1/F2} > ca.$ 50, complete spectral information can be recovered as measured with *SSIM*'s that are referenced to a noiseless spectrum.

3.5 2D Experiments

^{87}Rb (I = 3/2) MQMAS NMR experiments on $RbNO_3$ were conducted to acquire several datasets with different number of scans and with different modifications to the MQMAS pulse sequence to change the $SNR_{in}^{F1/F2}$ in each case. A whole-echo style sequence is used with an added SPAM pulse for sensitivity enhancement;[73] it is important to note that sequences with better sensitivity are available,[74–76] but for this proof-of-concept study we want to demonstrate lower *SNR* regimes. Experiments were carried out at 14.1 T with a spinning speed of 10 kHz. Datasets are the same size as the simulated cases with 1024 × 128 points ($t_2 \times t_1$) for the time domain and 4096 × 512 ($F_2 \times F_1$) for the frequency domain. The first example shows the 2D MQMAS-SPAM spectra acquired with 96 scans (**Figure 9a**). This is the minimum number of scans for the phase cycling, so this represents the lowest SNR_{in} for this type of acquisition. The 2D spectrum can be denoised in the same manner as described above for the simulated cases. The resulting SNR_{WT}'s are higher for both dimensions and the F_1 slices have increased $SSIM_{WT}$'s in comparison to the input data (**Figure 9b** and **Table 9**), where the *SSIM*'s are referenced to the spectra acquired with 382 scans, where SSIM(NS = 382, NS = 382) = 1.0000. A simple way to lower the experimental SNR_{in} is to use a less efficient MQMAS sequence, which was done by using weaker RF amplitudes for TQ-excitation and conversion pulses (**Table 9**) or by removing the SPAM pulse (**Figure 9c** and **Table 9**). The resulting SNR_{in}'s in the latter case are

approximately the lowest obtained for these experiments (*i.e.,* the experiment with 75 kHz RF fields features similar SNR_{in}'s). Nonetheless, the denoising helps increase the SNR_{WT} in both dimensions and increase the $SSIM_{WT}$ of all the extracted spectra (**Figure 9d** and **Table 9**). In both examples above it is important to note the reduction of t_1-noise, or an increase of SNR^{F1}. In experimental cases, t_1-noise originates from a combination of thermal noise and small spinning instabilities causing rotor desynchronization in the pulse sequence.[74] This can readily be observed as some of the more systematic t_1-noise forms ridges with well-defined slopes of F_2/F_1= +7/9, as would be expected for residual anisotropy in spin-3/2 triple-quantum MQMAS.

A final test case featuring lower *SNR* than ^{87}Rb experiments is the ^{17}O triple-quantum (3Q) MQMAS of the metal organic framework α-$Mg_3(HCOO)_6$ (**Figure 10**). This data was previously acquired by Martins *et al.* at the National High Magnetic Field Laboratory (NHMFL) using the series-connected hybrid (SCH) magnet operating at 35.2 T.[77] The data was acquired with a shifted-echo sequence at a spinning speed of 18 kHz. Since the pulse sequence is rotor-synchronized, the F_1 window is limited to 18 kHz, which causes aliasing from the SSB signal. This is remedied with Q-shearing and spectral zero-filling to expand the F_1 window (isotropic shearing, Q-shearing, and F_1 expansion protocols are included in DESPERATE).[72] The high-magnetic field strength results in substantial narrowing of the ^{17}O central-transitions patterns, yielding 2D spectra with a combination of narrow and broad features, making this an excellent test case for WT denoising.

The ^{17}O 3QMAS spectra contain patterns corresponding to twelve unique oxygen sites, with six of the CT patterns overlapped and centered around *ca.* 220 ppm on F_2 and the other six CT patterns are overlapped and centered around *ca.* 270 ppm (**Figure 10**). The features above and below these shifts correspond to SSB signal (**Figure 10 -** marked with *). The signal and

noise are dispersed over a slope due to the shearing and F_1 expansion (*vide supra*). Therefore, the *SNR* is measured as a single value using baseline noise isolated from this slope. $SNR_{in} = 60$ prior to any denoising, where minimal apodization is used (**Figure 10a**) and $SNR_{WT} = 589$ after WT denoising with $k = 3$ (**Figure 10b**). WT denoising also outperforms PCA denoising (**Figure S10**) where efficient denoising is challenging while also maintaining the resolution of the broad and narrow spectral features in the spectrum, similar to simulated MQMAS benchmarks (*c.f.* **Figure 7c**). The CT and SSB signal is clearly identifiable after denoising, which can allow for easier characterization of the system.

4. Conclusions

WT denoising is a robust and efficient technique for denoising 1D and 2D NMR data. Simulated and experimental results display improved performance over established SVD-based approaches for denoising. Frequency-domain WT denoising is reliable for use with different types of NMR spectra, including anisotropic dispersion-type spectra or isotropic high-resolution spectra, or combinations of both, as demonstrated with anisotropic powder patterns, CPMG spikelets, SSB manifolds, and 2D MQMAS spectra. WT denoising can be used on any type of NMR spectrum as evidenced by the variety in SSNMR spectra shown throughout (*i.e.,* different peak and pattern shapes). Large enhancements in *SNR* are somewhat trivial to obtain, whereas a key factor throughout this work is enhancements in *SSIMs*, which are indicative of partial or significant recovery of patterns or improvements in spectral uniformity in most cases. WT denoising is ultimately limited by the *SNR* of the input data, where there are certain limits at which a pattern can no longer be recovered/improved by denoising; in this work we find this limit to be $SNR \lesssim 50$ (*N.B.*: this is based on how we have measured *SNR* herein, which is not

consistent across the literature and in various software packages). Combined improvements in *SNR* and *SSIM* from denoising can allow for faster experimental acquisitions by necessitating less signal averaging and can be useful for improving the quality of regression analysis of NMR data (*e.g.,* fitting tensor parameters, relaxometry, *etc.*) where SNR is a limiting factor. The low computational cost of WT denoising as applied to 1D and 2D NMR datasets should make it an attractive option for many types of NMR experiments, especially for those involving the acquisition of CPMG and high-resolution datasets with large numbers of data points. Since the frequency-domain signal windowing for thresholding samples noise in both F_2 and F_1 dimensions, systematic t_1-noise (*i.e.,* not random thermal noise) is substantially reduced with WT denoising. SVD-based and WT denoising have been implemented in a free and open-source Python library called DESPERATE, which is designed for easy use by end users. We anticipate the aforementioned denoising techniques will find routine use in *n*D NMR and will be commonplace in the same way standard apodization is used.

5. Supplementary Material

See supplementary material for additional simulations, experiments, and experimental details.

6. Acknowledgments

R.W.S. and A.R.A would like to thank the National Science Foundation Chemical Measurement and Imaging Program, with partial co-funding from the Solid State and Materials Chemistry Program (NSF-2003854), for supporting this work. The National High Magnetic Field Laboratory is supported by the National Science Foundation through NSF/DMR-1644779 and

the State of Florida. Los Alamos National Laboratory is operated by Triad National Security, LLC, for the National Nuclear Security Administration of US Department of Energy (contract no. 89233218CNA000001). Yining Huang, Zhehong Gan, and Ivan Hung are thanked for providing the experimental ^{17}O MQMAS dataset and advice.

7. Data Availability

All simulation input files, processing functions and scripts, and NMR datasets are available on github.com/rschurko/desperate.

8. References

(1) Fu, R.; Brey, W. W.; Shetty, K.; Gor'kov, P.; Saha, S.; Long, J. R.; Grant, S. C.; Chekmenev, E. Y.; Hu, J.; Gan, Z.; et al. *J. Magn. Reson.* **2005**, *177*, 1–8.

(2) Gan, Z.; Hung, I.; Wang, X.; Paulino, J.; Wu, G.; Litvak, I. M.; Gor'kov, P. L.; Brey, W. W.; Lendi, P.; Schiano, J. L.; et al. *J. Magn. Reson.* **2017**, *284*, 125–136.

(3) Matsuki, Y.; Nakamura, S.; Hobo, F.; Endo, Y.; Takahashi, H.; Suematsu, H.; Fujiwara, T. *J. Magn. Reson.* **2022**, *335*, 107139.

(4) Maly, T.; Debelouchina, G. T.; Bajaj, V. S.; Hu, K.-N.; Joo, C.-G.; Mak–Jurkauskas, M. L.; Sirigiri, J. R.; van der Wel, P. C. A.; Herzfeld, J.; Temkin, R. J.; et al. *J. Chem. Phys.* **2008**, *128*, 052211.

(5) Ardenkjaer-Larsen, J. H. *J. Magn. Reson.* **2016**, *264*, 3–12.

(6) Rossini, A. J.; Zagdoun, A.; Lelli, M.; Lesage, A.; Copéret, C.; Emsley, L. *Acc. Chem. Res.* **2013**, *46*, 1942–1951.

(7) Neuhaus, D. *Nuclear Overhauser Effect*. In *Encyclopedia of Magnetic Resonance*; John Wiley & Sons, Ltd: Chichester, UK, 2011.

(8) Hartmann, S. R.; Hahn, E. L. *Phys. Rev.* **1962**, *128*, 2042–2053.

(9) Pines, A.; Gibby, M. G.; Waugh, J. S. *J. Chem. Phys.* **1972**, *56*, 1776–1777.

(10) Pines, A.; Gibby, M. G.; Waugh, J. S. *J. Chem. Phys.* **1973**, *59*, 569–590.

(11) Wolf, T.; Jaroszewicz, M. J.; Frydman, L. *J. Phys. Chem. C* **2021**, *125*, 1544–1556.

(12) Carr, H. Y.; Purcell, E. M. *Phys. Rev.* **1954**, *94*, 630–638.

(13) Meiboom, S.; Gill, D. *Rev. Sci. Instrum.* **1958**, *29*, 688–691.

(14) Otting, G. *Annu. Rev. Biophys.* **2010**, *39*, 387–405.

(15) Cai, S.; Seu, C.; Kovacs, Z.; Sherry, A. D.; Chen, Y. *J. Am. Chem. Soc.* **2006**, *128*, 13474–13478.

(16) Mueller, L. *J. Am. Chem. Soc.* **1979**, *101*, 4481–4484.

(17) Bodenhausen, G.; Ruben, D. J. *Chem. Phys. Lett.* **1980**, *69*, 185–189.

(18) Bax, A.; Griffey, R. H.; Hawkins, B. L. *J. Magn. Reson.* **1983**, *55*, 301–315.

(19) Jaroszewicz, M. J.; Altenhof, A. R.; Schurko, R. W.; Frydman, L. *J. Am. Chem. Soc.* **2021**, *143*, 19778–19784.

(20) Lee, J. H.; Okuno, Y.; Cavagnero, S. *J. Magn. Reson.* **2014**, *241*, 18–31.

(21) Lesage, A. *Phys. Chem. Chem. Phys.* **2009**, *11*, 6876.

(22) Bartholdi, E.; Ernst, R. R. *J. Magn. Reson.* **1973**, *11*, 9–19.

(23) Levitt, M. H. *Spin Dynamics*; John Wiley & Sons Ltd, 2008.

(24) Golyandina, N.; Nekrutkin, V.; Zhigljavsky, A. A. *Analysis of time series structure: SSA and related techniques*; CRC Press, 2001.

(25) Chiron, L.; van Agthoven, M. A.; Kieffer, B.; Rolando, C.; Delsuc, M.-A. *Proc. Natl. Acad. Sci.* **2014**, *111*, 1385–1390.

(26) Gauthier, J. R.; Mabury, S. A. *Anal. Chem.* **2022**, acs.analchem.1c05107.

(27) Koprivica, D.; Martinho, R. P.; Novakovic, M.; Michael, J.; Frydman, L. *J. Magn. Reson.*

2022, 107187.

(28) Cadzow, J. A. *IEEE Trans. Acoust.* **1988**, *36*, 49–62.

(29) Brissac, C.; Malliavin, T.; Delsuc, M. *J. Biomol. NMR* **1995**, *6*, 361–365.

(30) Man, P. P.; Bonhomme, C.; Babonneau, F. *Solid State Nucl. Magn. Reson.* **2014**, *61–62*, 28–34.

(31) Laurent, G.; Woelffel, W.; Barret-Vivin, V.; Gouillart, E.; Bonhomme, C. *Appl. Spectrosc. Rev.* **2019**, *54*, 602–630.

(32) Laurent, G.; Gilles, P.-A.; Woelffel, W.; Barret-Vivin, V.; Gouillart, E.; Bonhomme, C. *Appl. Spectrosc. Rev.* **2020**, *55*, 173–196.

(33) Kusaka, Y.; Hasegawa, T.; Kaji, H. *J. Phys. Chem. A* **2019**, *123*, 10333–10338.

(34) Petrov, O. V.; Lang, J.; Vogel, M. *J. Magn. Reson.* **2021**, *326*, 106965.

(35) Mason, H. E.; Uribe, E. C.; Shusterman, J. A. *Phys. Chem. Chem. Phys.* **2018**, *20*, 18082–18088.

(36) Altenhof, A. R.; Jaroszewicz, M. J.; Frydman, L.; Schurko, R. *Phys. Chem. Chem. Phys.* **2022**.

(37) Leung, A. K. man; Chau, F. tim; Gao, J. bin. *Chemom. Intell. Lab. Syst.* **1998**, *43*, 165–184.

(38) Shao, X.-G.; Leung, A. K.-M.; Chau, F.-T. *Acc. Chem. Res.* **2003**, *36*, 276–283.

(39) Donoho, D. L.; Johnstone, I. M. *Biometrika* **1994**, *81*, 425–455.

(40) Donoho, D. L. *IEEE Trans. Inf. Theory* **1995**, *41*, 613–627.

(41) Bruce, A. G.; Gao, H.-Y. *WaveShrink: Shrinkage Functions and Thresholds*. In *Wavelet Applications in Signal and Image Processing III*; Laine, A. F., Unser, M. A., Eds.; 1995; Vol. 2569, pp 270–281.

(42) Gao, H. Y.; Bruce, A. G. *Stat. Sin.* **1997**, *7*, 855–874.

(43) Breiman, L. *Technometrics* **1995**, *37*, 373–384.

(44) Gao, H. Y. *J. Comput. Graph. Stat.* **1998**, *7*, 469–488.

(45) Srivastava, M.; Anderson, C. L.; Freed, J. H. *IEEE Access* **2016**, *4*, 3862–3877.

(46) Srivastava, M.; Georgieva, E. R.; Freed, J. H. *J. Phys. Chem. A* **2017**, *121*, 2452–2465.

(47) Srivastava, M.; Dzikovski, B.; Freed, J. H. *J. Phys. Chem. A* **2021**, *125*, 4480–4487.

(48) Ge, X.; Fan, Y.; Li, J.; Wang, Y.; Deng, S. *J. Magn. Reson.* **2015**, *251*, 71–83.

(49) Monaretto, T.; Souza, A.; Moraes, T. B.; Bertucci-Neto, V.; Rondeau-Mouro, C.; Colnago, L. A. *Magn. Reson. Chem.* **2019**, *57*, 616–625.

(50) Günther, U. L.; Ludwig, C.; Rüterjans, H. *J. Magn. Reson.* **2000**, *145*, 201–208.

(51) Zhong, J.; Sleighter, R. L.; Salmon, E.; McKee, G. A.; Hatcher, P. G. *Org. Geochem.* **2011**, *42*, 903–916.

(52) Frydman, L.; Harwood, J. S. *J. Am. Chem. Soc.* **1995**, *117*, 5367–5368.

(53) Medek, A.; Harwood, J. S.; Frydman, L. *J. Am. Chem. Soc.* **1995**, *117*, 12779–12787.

(54) Altenhof, A. R.; Lindquist, A. W.; Foster, L. D. D.; Holmes, S. T.; Schurko, R. W. *J. Magn. Reson.* **2019**, *309*, 106612.

(55) Harris, K. J.; Lupulescu, A.; Lucier, B. E. G.; Frydman, L.; Schurko, R. W. *J. Magn. Reson.* **2012**, *224*, 38–47.

(56) Massiot, D.; Touzo, B.; Trumeau, D.; Coutures, J. P.; Virlet, J.; Florian, P.; Grandinetti, P. J. *Solid State Nucl. Magn. Reson.* **1996**, *6*, 73–83.

(57) Gan, Z.; Grant, D. M.; Ernst, R. R. *Chem. Phys. Lett.* **1996**, *254*, 349–357.

(58) Bak, M.; Rasmussen, J. T.; Nielsen, N. C. *J. Magn. Reson.* **2011**, *213*, 366–400.

(59) Juhl, D. W.; Tošner, Z.; Vosegaard, T. *Annu. Reports NMR Spectrosc.* **2020**, 1–59.

(60) Tošner, Z.; Andersen, R.; Stevensson, B.; Edén, M.; Nielsen, N. C.; Vosegaard, T. *J. Magn. Reson.* **2014**, *246*, 79–93.

(61) Edén, M. *Concepts Magn. Reson. Part A Bridg. Educ. Res.* **2003**, *18*, 24–55.

(62) Brown, S. P.; Wimperis, S. *J. Magn. Reson.* **1997**, *128*, 42–61.

(63) Harris, C. R.; Millman, K. J.; van der Walt, S. J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N. J.; et al. *Nature* **2020**, *585*, 357–362.

(64) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. *Nat. Methods* **2020**, *17*, 261–272.

(65) Lee, G.; Gommers, R.; Waselewski, F.; Wohlfahrt, K.; O'Leary, A. *J. Open Source Softw.* **2019**, *4*, 1237.

(66) Hunter, J. D. *Comput. Sci. Eng.* **2007**, *9*, 90–95.

(67) Wang, X.; Dai, Y. *An Improved Denoising Method Based on Stationary Wavelet*

Transform. In *Proceedings of the 2018 International Symposium on Communication Engineering & Computer Science (CECS 2018)*; Atlantis Press: Paris, France, 2018; Vol. 86, pp 481–485.

(68) Wang, Z.; Bovik, A. C.; Sheikh, H. R.; Simoncelli, E. P. *IEEE Trans. Image Process.* **2004**, *13*, 600–612.

(69) O'Dell, L. A.; Schurko, R. W. *Chem. Phys. Lett.* **2008**, *464*, 97–102.

(70) Koppe, J.; Bußkamp, M.; Hansen, M. R. *J. Phys. Chem. A* **2021**, acs.jpca.1c02958.

(71) Altenhof, A. R.; Jaroszewicz, M. J.; Lindquist, A. W.; Foster, L. D. D.; Veinberg, S. L.; Schurko, R. W. *J. Phys. Chem. C* **2020**, *124*, 14730–14744.

(72) Hung, I.; Trébosc, J.; Hoatson, G. L.; Vold, R. L.; Amoureux, J. P.; Gan, Z. *J. Magn. Reson.* **2009**, *201*, 81–86.

(73) Gan, Z.; Kwak, H. *J. Magn. Reson.* **2004**, *168*, 346–351.

(74) Hung, I.; Gan, Z. *J. Magn. Reson.* **2021**, *328*, 106994.

(75) Hung, I.; Gan, Z. *J. Magn. Reson.* **2021**, *324*, 106913.

(76) Koppe, J.; Knitsch, R.; Wegner, S.; Hansen, M. R. *J. Magn. Reson.* **2020**, *321*, 106873.

(77) Martins, V.; Xu, J.; Wang, X.; Chen, K.; Hung, I.; Gan, Z.; Gervais, C.; Bonhomme, C.; Jiang, S.; Zheng, A.; et al. *J. Am. Chem. Soc.* **2020**, *142*, 14877–14889.

Figure Captions

Scheme 1: Flowchart of the proposed wavelet transform (WT) denoising routine. Each step is described in detail in the main text. A_k and D_k are the approximation and detail components, respectively. SWT and ISWT are the undecimated stationary wavelet transform and inverse-SWT, respectively.

Figure 1. Simulated ^{119}Sn CPMG spectra that are processed with echo coaddition and added Gaussian noise (left column), which are then separately denoised with Cadzow (middle column) or the WT routine (right column). (a-h) The amount of added Gaussian noise increases in every column from top to bottom, resulting in progressively decreasing peak-to-peak SNR (SNR_p) for the input spectra. Resulting SNR_p's and *SSIM*'s of the input and denoised data are summarized in **Table 1**. All spectra are scaled to the same absolute intensity for visual clarity.

Figure 2. Simulated ^{119}Sn CPMG spikelet spectra that are processed without echo coaddition and added Gaussian noise (left column) and then denoised with the WT routine (right column). (a-h) The amount of added Gaussian noise increases from top to bottom in each column, resulting in progressively decreasing SNR_p. The inset shows the *ca.* –82 to –92 kHz frequency region of the CPMG spectrum in (e), with spikelets from the input (top) and denoised (bottom) spectrum compared for visual clarity. Resulting SNR_p's and *SSIM*'s of the input and denoised spectra are summarized in **Table 2**. All spectra are scaled to the same absolute intensity for visual clarity. These spectra are simulated with a smaller span in comparison to those shown in **Figure 1**, in order to resolve the spikelets for visual clarity (see text).

Figure 3. Experimental ^{119}Sn WURST-CPMG spectra of SnO that are processed with echo coaddition (left column), which are then separately denoised with Cadzow (middle column) or the WT routine (right column). (a-h) The number of experimental scans (NS) in every column decreases from top to bottom, resulting in progressively decreasing SNR_p for the input spectra. Resulting SNR_p's and *SSIM*'s of the input and denoised spectra are summarized in **Table 3**. All spectra are scaled to the same absolute intensity for visual clarity.

Figure 4. Experimental ^{119}Sn WURST-CPMG spikelet spectra of SnO that are processed without echo coaddition (left column) and then denoised with the WT routine (right column). (a-h) The number of experimental scans (NS) in every column decreases from top to bottom, resulting in progressively decreasing SNR_p for the input spectra. The inset shows the *ca.* –175 to –195 kHz frequency region of the CPMG spectrum in (e) for visual clarity of input (top) and denoised (bottom) spikelets. Resulting SNR_p's and *SSIM*'s of the input and denoised data are summarized in **Table 4**. All spectra are scaled to the same absolute intensity for visual clarity.

Figure 5. Experimental ^{195}Pt WURST-CPMG/MAS spectra of $Pt(NH_3)_4Cl_2 \cdot H_2O$ that are processed with echo coaddition (left column), then separately denoised with Cadzow (middle column) or the WT routine (right column). (a-g) The number of experimental scans (NS) in every column decreases from top to bottom, resulting in progressively decreasing SNR_p for the input spectra. Resulting SNR_p's and *SSIM*'s of the input and denoised data are summarized in **Table 5**. All spectra are scaled to the same absolute intensity for visual clarity.

Figure 6. Experimental ^{23}Na Bloch decay spectra of a mixture of NaCl:Na_2SO_4 3:50 w/w mixture that are processed (left column), which are then separately denoised with Cadzow (middle column) or the WT routine (right column). (a-h) The number of experimental scans (NS) in every column decreases from top to bottom, resulting in progressively decreasing SNR_p for the input spectra. Resulting SNR_p's and *SSIM*'s of the input and denoised data are summarized in **Table 6**. The *SSIM*'s are referenced to an experimental spectrum acquired with 256 scans, which is not shown for brevity. All spectra are scaled to the same absolute intensity for visual clarity.

Figure 7. Simulated ^{87}Rb MQMAS spectra (a) with added Gaussian noise, and then denoised with either (b) incremented Cadzow over the t_1 dimension; (c) PCA on the entire spectral matrix; or (d) the 2D WT routine. See **Table 7** for *SNR* measurements. The F_1 projection is represented as a sum of all F_2 slices, while the F_2 projection is represented as a maximum intensity skyline over all F_1 slices (see supporting information for comparisons of projections). The F_1 projections are shown to the right of each contour plot and represent the separated patterns associated with each ^{87}Rb site. MQMAS datasets are simulated in SIMPSON in the time domain using a three-pulse whole-echo sequence with ideal pulses (*i.e.*, infinite power and no finite pulse widths) that have optimized tip angles for TQC excitation and conversion. The baseline contour is set to 2% of the maximum spectral intensity and 30 contours are drawn in every case. All spectral matrices are scaled to the same absolute intensity for visual clarity.

Figure 8. Simulated ^{87}Rb MQMAS spectra (a,c) with added Gaussian noise, and then denoised with (b,d) the 2D WT routine. The $SNR_{p,in}$ over F_1 and F_2 is higher in (a) than in (c). See **Table 8** for all metrics. The F_1 projection is represented as a sum of all F_2 slices, while the F_2 projection is represented as a maximum intensity skyline over all F_1 slices (see supporting information for comparisons of projections). The F_1 projections are shown to the right of each contour plot and represent the separated patterns associated with each ^{87}Rb site. The baseline contour is set to 2% of the maximum spectral intensity in every case and 30 contours are drawn. All spectral matrices are scaled to the same absolute intensity for visual clarity.

Figure 9. Experimental ^{87}Rb MQMAS spectra (a) acquired with a SPAM pulse and 96 scans and then (b) WT denoised, or (c) acquired without a SPAM pulse and 96 scans and then (d) WT denoised. See **Table 9** for all metrics. The F_1 projection is represented as a sum of all F_2 slices, while and the F_2 projection is represented as a maximum intensity skyline over all F_1 slices (see supporting information for comparisons of projections). The F_1 projections are shown to the right of each contour plot and represent the separated patterns associated with each ^{87}Rb site. The baseline contour is set to 0.5% of the maximum spectral intensity in every case and 40 contours are drawn. All spectral matrices are scaled to the same absolute intensity for visual clarity.

Figure 10. Experimental ^{17}O MQMAS spectra of α-$Mg_3(HCOO)_6$ (a) acquired with a shifted-echo sequence at 35.2 T with a spinning speed of 18 kHz and then (b) WT denoised. The spectra are processed by first Q-shearing with a shearing constant, $k_q = 3$ for 3QMAS, zero-filling the F_1 spectral window to expand the aliased signal of the SSBs, reverse Q-shearing, and then isotropic shearing with $k_q = 19/12$. Signal associated with SSBs are marked by red asterisks (*). The F_1 projection is represented as a sum of all F_2 slices, while the F_2 projection is represented as a maximum intensity skyline over all F_1 slices. The baseline contour is set to 2% of the maximum spectral intensity in each case and 30 contours are drawn.

Table 1: SNR_p's and *SSIM*'s for spectra in **Figure 1** prior to denoising, after Cadzow denoising, and after WT denoising.

Spectrum	$SNR_{p,in}$	$SNR_{p,C}$	$SNR_{p,WT}$	$SSIM_{in}$	$SSIM_{C}$	$SSIM_{WT}$
a	452	16310	14669	0.9997	0.9996	0.9999
b	248	26980	5353	0.9991	0.9994	0.9996
c	171	18212	3756	0.9975	0.9985	0.9990
d	105	17375	3315	0.9957	0.9974	0.9984
e	723	29320	1185	0.9934	0.9970	0.9978
f	64	24127	2233	0.9882	0.9660	0.9862
g	57	17653	2595	0.9809	0.9666	0.9807
h	29	9488	991	0.9636	0.9305	0.8485

Subscripts 'in', 'C', and 'WT' refer to the input, Cadzow denoised, and WT denoised spectra, respectively.

Table 2: SNR_p's and *SSIM*'s for spectra in **Figure 2** prior to denoising and after WT denoising.

Spectrum	$SNR_{p,in}$	$SNR_{p,WT}$	$SSIM_{in}$	$SSIM_{WT}$	$SSIM_{in,low}$	$SSIM_{WT,low}$
a	410	15054	0.9409	0.8769	0.8355	0.8179
b	274	10407	0.8766	0.8483	0.6935	0.8166
c	206	7638	0.8005	0.8312	0.5632	0.8122
d	138	5655	0.6402	0.8189	0.3648	0.8022
e	105	4023	0.4998	0.8069	0.2443	0.7688
f	70	2673	0.3070	0.7639	0.1213	0.6527
g	50	2056	0.2000	0.7013	0.0734	0.4962
h	41	1653	0.1375	0.6497	0.0474	0.3470

The subscript "low" indicates that the *SSIM* is measured only over the low-frequency half of the spectrum that spans *ca*. –30 to – 90 kHz.
Subscripts 'in', 'C', and 'WT' refer to the input, Cadzow denoised, and WT denoised spectra, respectively.

Table 3: SNR_p's and *SSIM*'s for spectra in **Figure 3** prior to denoising, after Cadzow denoising, and after WT denoising.

Spectrum (scans)	$SNR_{p,in}$	$SNR_{p,C}$	$SNR_{p,WT}$	$SSIM_{in}$	$SSIM_{C}$	$SSIM_{WT}$
a (1024)	494	1591	14529	1.0000	0.9985	0.9999
b (512)	415	1445	7679	0.9998	0.9984	0.9999
c (256)	275	2259	5180	0.9996	0.9991	0.9998
d (128)	180	1448	9605	0.9992	0.9983	0.9998
e (64)	158	1693	7014	0.9989	0.9987	0.9997
f (32)	85	2173	1664	0.9966	0.9977	0.9978
g (16)	58	1418	3329	0.9903	0.9931	0.9919
h (8)	40	1011	708	0.9878	0.9912	0.9974

Subscripts 'in', 'C', and 'WT' refer to the input, Cadzow denoised, and WT denoised spectra, respectively.

Table 4: SNR_p's and *SSIM*'s for spectra in **Figure 4** prior to denoising and after WT denoising.

Spectrum (scans)	$SNR_{p,in}$	$SNR_{p,WT}$	$SSIM_{in}$	$SSIM_{WT}$
a (1024)	308	1788	1.0000	0.9750
b (512)	236	2732	0.9893	0.9543
c (256)	182	2474	0.9784	0.9270
d (128)	129	4079	0.9657	0.8868
e (64)	97	3144	0.9484	0.8481
f (32)	72	2140	0.8992	0.8147
g (16)	52	1636	0.8335	0.7441
h (8)	35	1197	0.7209	0.6811

Subscripts 'in', 'C', and 'WT' refer to the input, Cadzow denoised, and WT denoised spectra, respectively.

Table 5: SNR_p's and *SSIM*'s for spectra in **Figure 5** prior to denoising, after Cadzow denoising, and after WT denoising.

Spectrum	$SNR_{p,in}$	$SNR_{p,C}$	$SNR_{p,WT}$	$SSIM_{in}$	$SSIM_C$	$SSIM_{WT}$
a (2048)	166	1816016	1943	1.0000	0.9636	0.9983
b (1024)	124	603390	2534	0.9874	0.9901	0.9888
c (512)	79	676	1929	0.9781	0.9842	0.9800
d (256)	56	176	1122	0.9748	0.9864	0.9855
e (128)	25	73	357	0.7985	0.8551	0.8614
f (64)	21	69	260	0.8514	0.9173	0.9229

Subscripts 'in', 'C', and 'WT' refer to the input, Cadzow denoised, and WT denoised spectra, respectively.

Table 6: SNR_p's and *SSIM*'s for spectra in **Figure 6** prior to denoising, after Cadzow denoising, and after WT denoising.

Spectrum	$SNR_{p,in}$	$SNR_{p,C}$	$SNR_{p,WT}$	$SSIM_{in}$	$SSIM_C$	$SSIM_{WT}$
a (128)	299	704	623	1.0000	1.0000	1.0000
b (64)	226	730	563	0.9999	1.0000	0.9999
c (32)	171	524	517	0.9998	0.9999	0.9999
d (16)	124	525	435	0.9997	0.9999	0.9999
e (8)	94	558	324	0.9992	0.9998	0.9999
f (4)	61	228	292	0.9984	0.9997	0.9995
g (2)	46	133	195	0.9970	0.9992	0.9993
h (1)	33	264	67	0.9941	0.9982	0.9974

Subscripts 'in', 'C', and 'WT' refer to the input, Cadzow denoised, and WT denoised spectra, respectively.

Table 7: *SNR*'s and *SSIM*'s for spectra in **Figure 7** prior to denoising, after Cadzow denoising and after WT denoising.

SNR_{in}^{F1}	SNR_{C}^{F1}	SNR_{PCA}^{F1}	SNR_{WT}^{F1}	SNR_{in}^{F2}	SNR_{C}^{F2}	SNR_{PCA}^{F2}	SNR_{WT}^{F2}
102	102	168	280	81	273	64	697

Superscripts F1 and F2 denote measurements of the *SNR* from the indirect and direct frequency dimensions, F_1 and F_2, respectively.
Subscripts 'in', 'C', 'PCA', and 'WT' refer to the input, Cadzow denoised, PCA denoised, and WT denoised spectra, respectively.

Table 8: *SNR*'s and *SSIM*'s for spectra in **Figure 7d** and **Figure 8** prior to denoising, after Cadzow denoising and after WT denoising.

Spec	SNR_{in}^{F1}	SNR_{WT}^{F1}	SNR_{in}^{F2}	SNR_{WT}^{F2}	$SSIM_{in}^{1}$	$SSIM_{WT}^{1}$	$SSIM_{in}^{2}$	$SSIM_{WT}^{2}$	$SSIM_{in}^{3}$	$SSIM_{WT}^{3}$
7d	102	280	81	697	0.9951	0.9963	0.9938	0.9945	0.9964	0.9989
8a,b	58	1284	50	430	0.9879	0.9922	0.9847	0.9839	0.9916	0.9959
8c,d	50	1150	48	408	0.9792	0.9893	0.9736	0.9820	0.9876	0.9875

The 'Spec' column refers to spectra shown in their corresponding figures (*e.g.,* 7d for **Figure 7d**). The unlabelled ones are not shown in any figure for brevity.
Superscripts F1 and F2 denote measurements of the *SNR* from the indirect and direct frequency dimensions, F_1 and F_2, respectively.
Subscripts 'in', and 'WT' refer to the input, and WT denoised spectra, respectively.
Superscripts 1, 2, and 3 indicate the F_1 slice spectra from top to bottom in MQMAS figures, respectively.

Table 9: *SNR*'s and *SSIM*'s for spectra in **Figure 9** prior to denoising and after WT denoising.

Spec	SNR_{in}^{F1}	SNR_{WT}^{F1}	SNR_{in}^{F2}	SNR_{WT}^{F2}	$SSIM_{in}^{1}$	$SSIM_{WT}^{1}$	$SSIM_{in}^{2}$	$SSIM_{WT}^{2}$	$SSIM_{in}^{3}$	$SSIM_{WT}^{3}$
384	255	434	1173	35438249	1.0000	0.9995	1.0000	0.9987	1.0000	0.9999
192	161	212	620	1275283	0.9858	0.9819	0.9829	0.9792	0.9544	0.9514
96 (8a,b)	130	205	559	884341	0.9791	0.9759	0.9679	0.9706	0.9327	0.9389
75 kHz	63	104	262	26032	0.9956	0.9972	0.9928	0.9963	0.9939	0.9958
No SPAM (8c,d)	80	118	245	28315	0.9915	0.9885	0.9821	0.9729	0.9963	0.9956

The first three rows in the 'Spec' column indicate the number of scans used in an experiment with a sequence that uses 100 kHz RF for TQ-excitation and conversion and a SPAM pulse. The fourth row uses only 75 kHz for TQ excitation and conversion (75 kHz) and 96 scans, and the last row uses no SPAM pulse with 100 kHz RF for TQ excitation and conversion and 96 scans. Corresponding figures are noted where applicable (*e.g.,* 8a,b for **Figure 9a,b**).
Superscripts F1 and F2 denote measurements of the *SNR* from the indirect and direct frequency dimensions, F_1 and F_2, respectively.
Subscripts 'in', and 'WT' refer to the input and WT spectra, respectively.
Superscripts 1, 2, and 3 indicate the F_1 slice spectra from top to bottom in MQMAS figures, respectively.

Declaration of interests

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

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Scheme 1

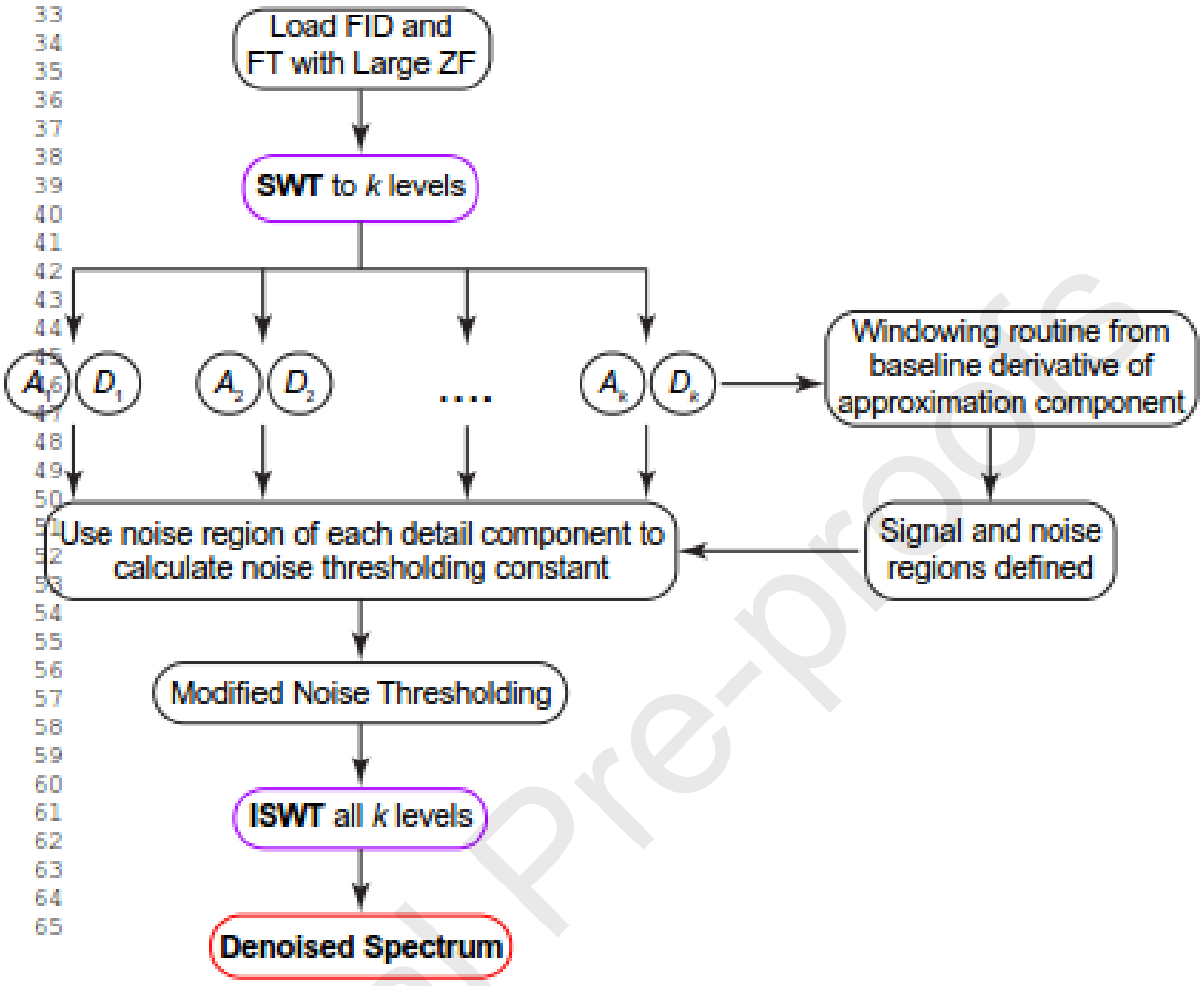


FIG. 1.

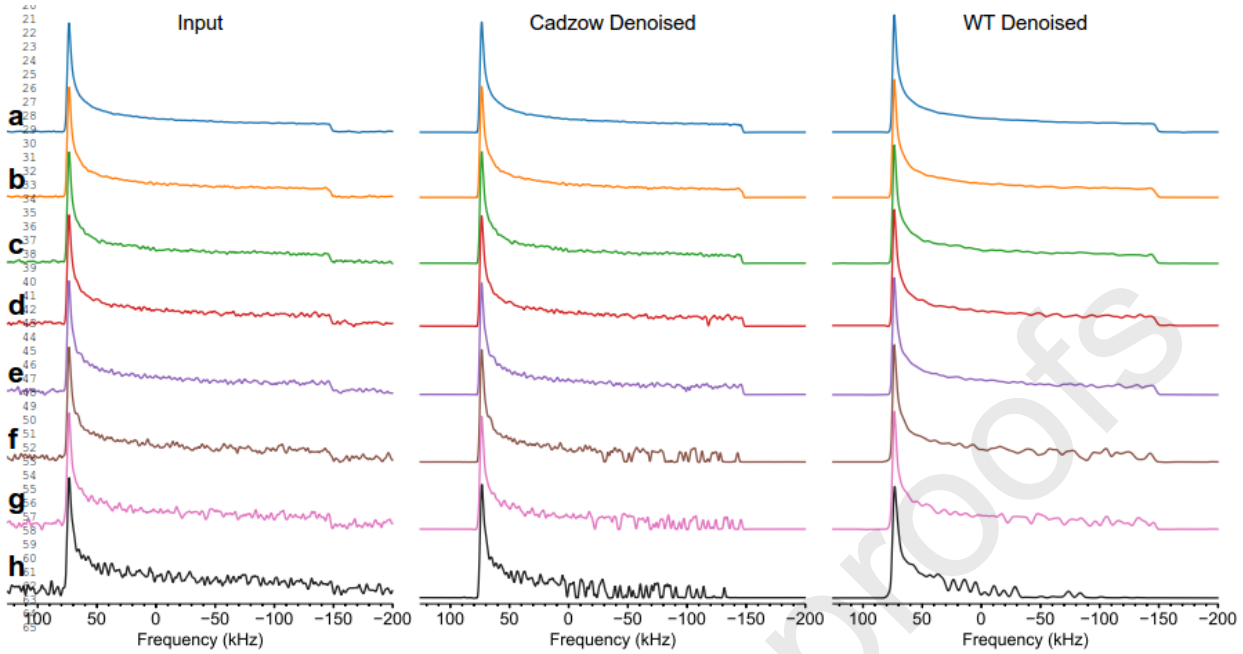

Input
Cadzow Denoised
WT Denoised
a
b
c
d
e
f
g
h
100 50 0 −50 −100 −150 −200
Frequency (kHz)

FIG. 2.

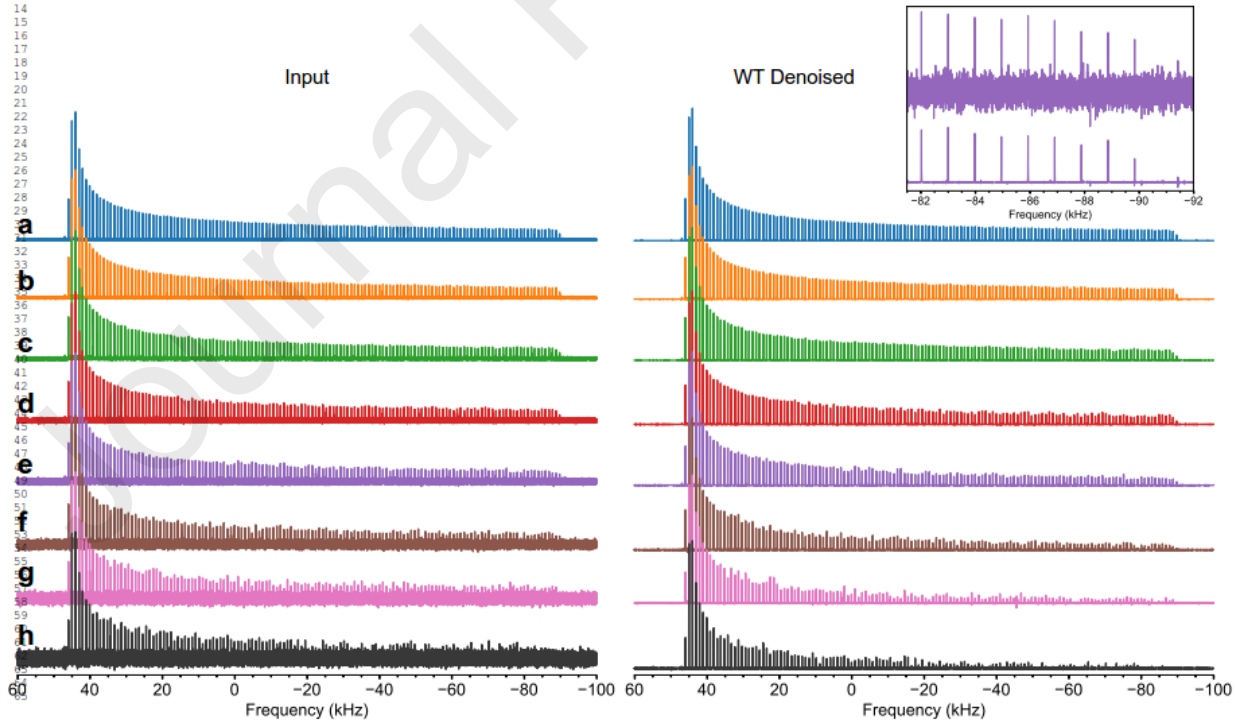

Input
WT Denoised
−82 −84 −86 −88 −90 −92
Frequency (kHz)
a
b
c
d
e
f
g
h
60 40 20 0 −20 −40 −60 −80 −100
Frequency (kHz)

FIG. 3.

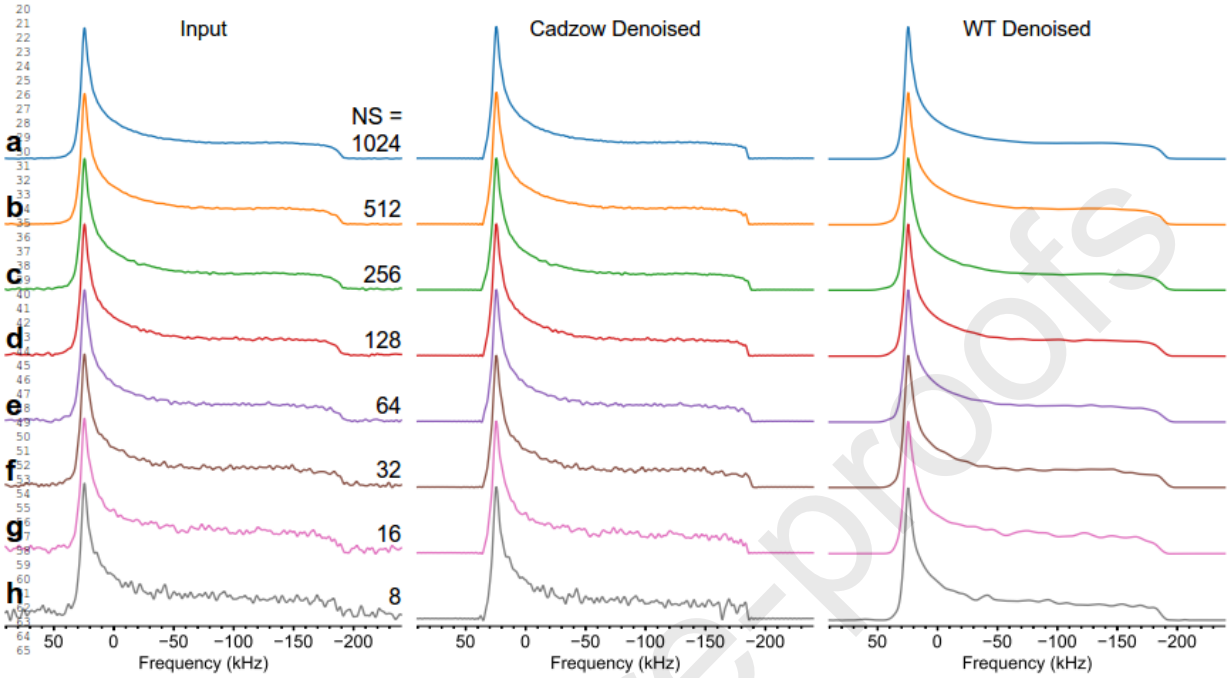


FIG. 4.

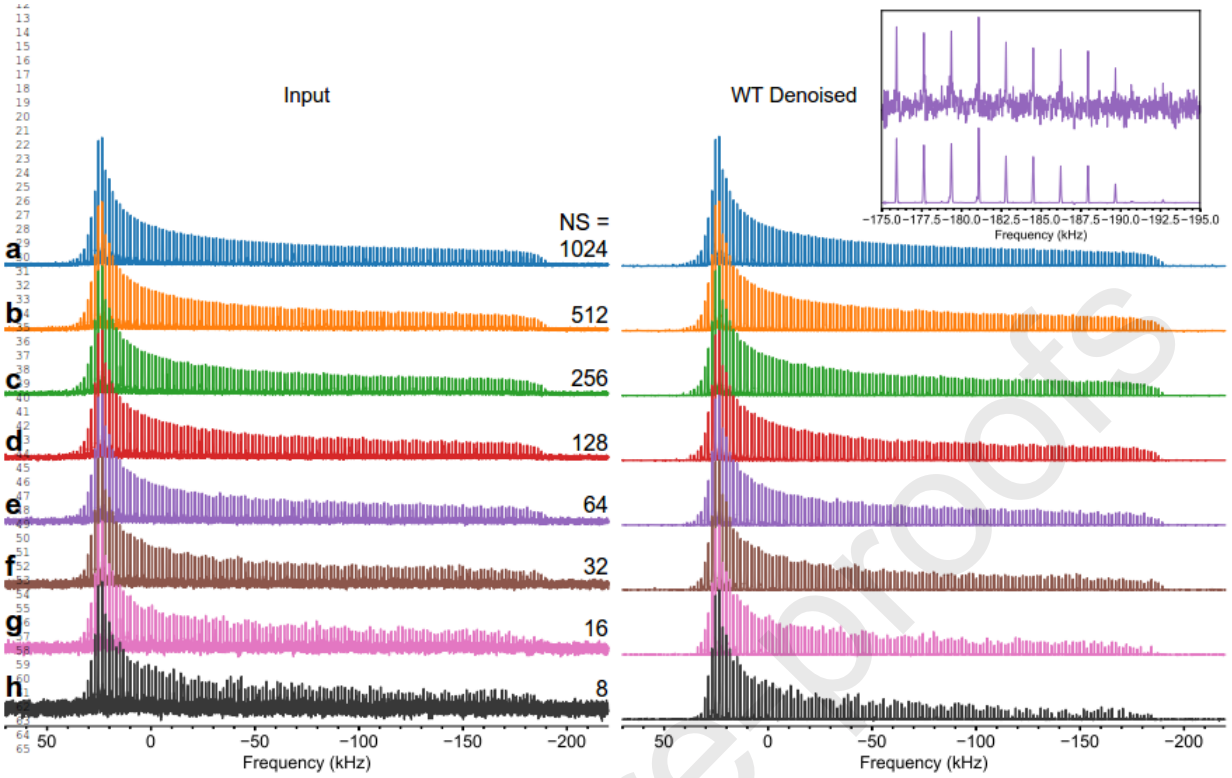


FIG. 5.

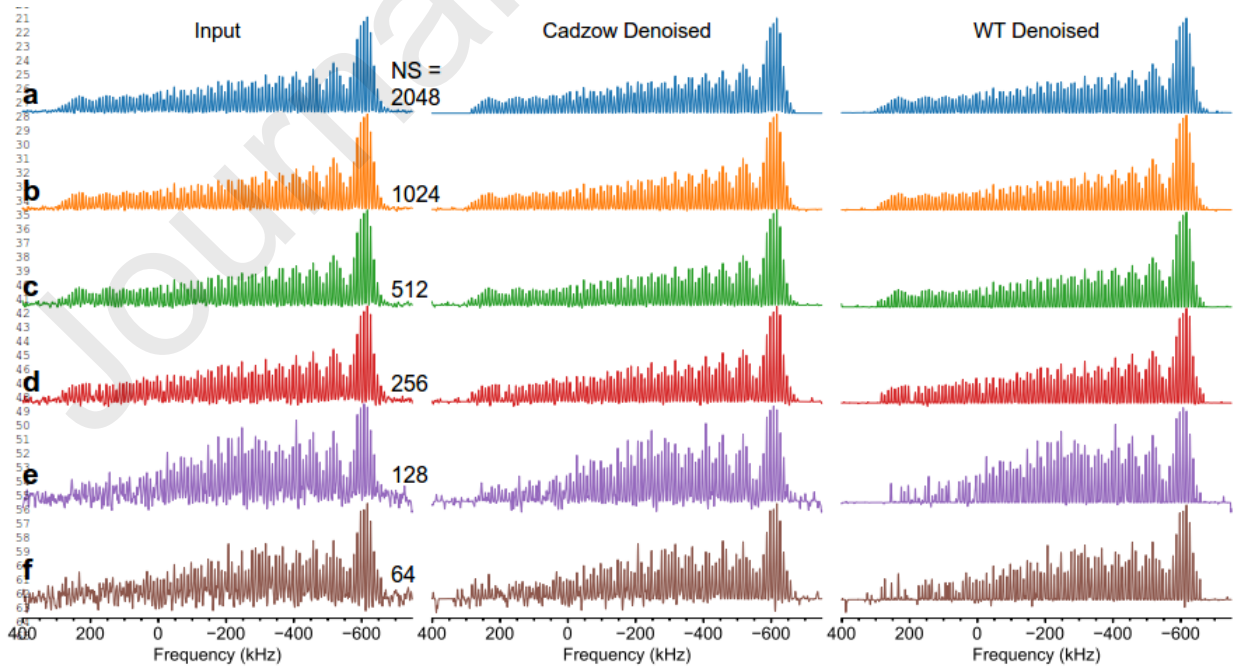


FIG. 6.

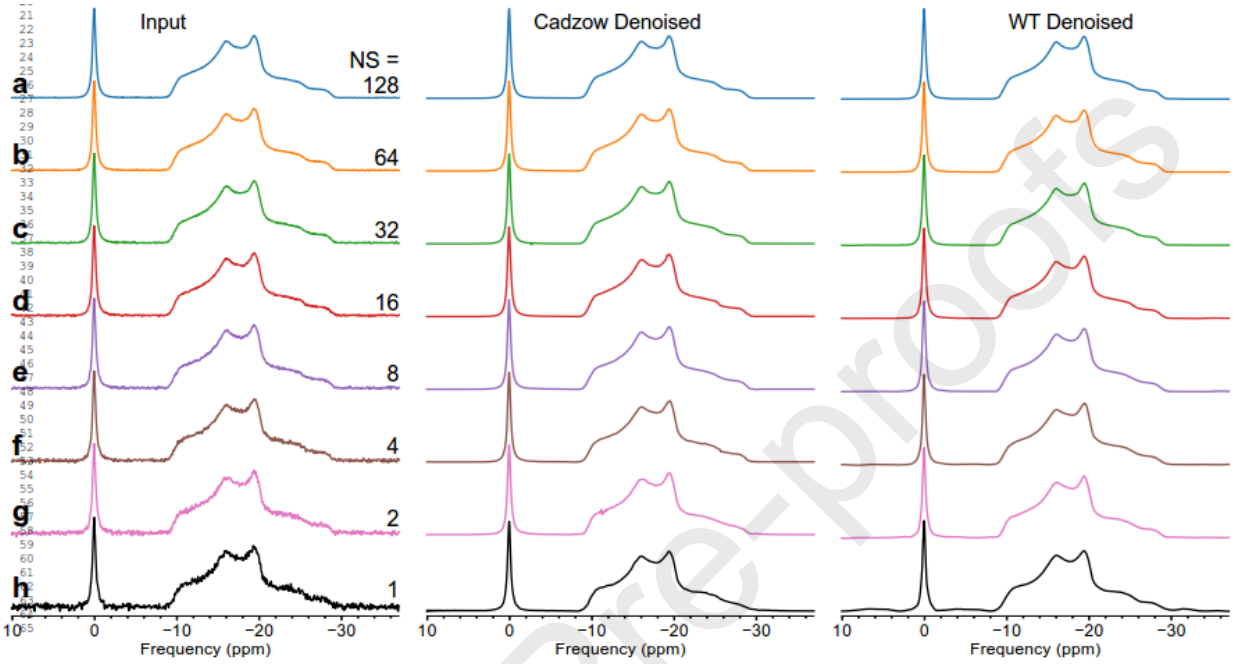


FIG. 7.

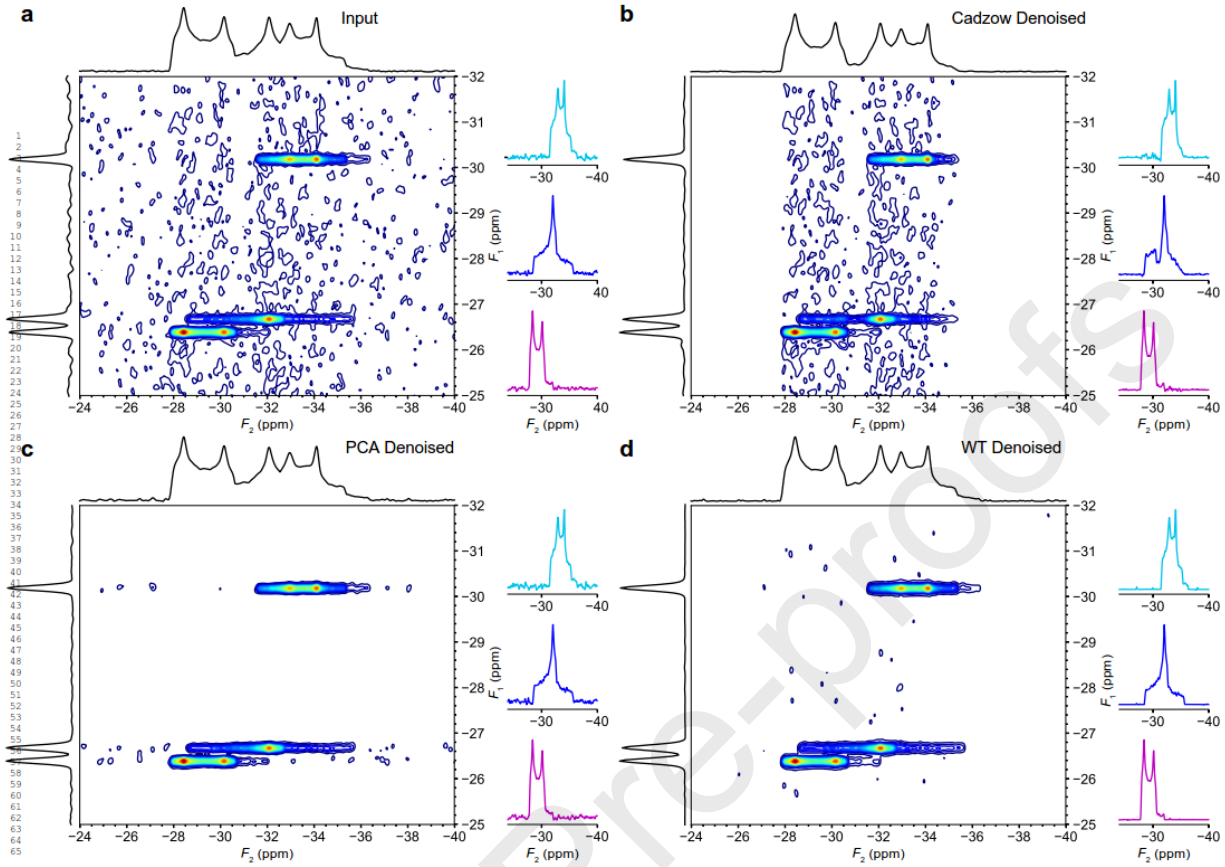


FIG. 8.

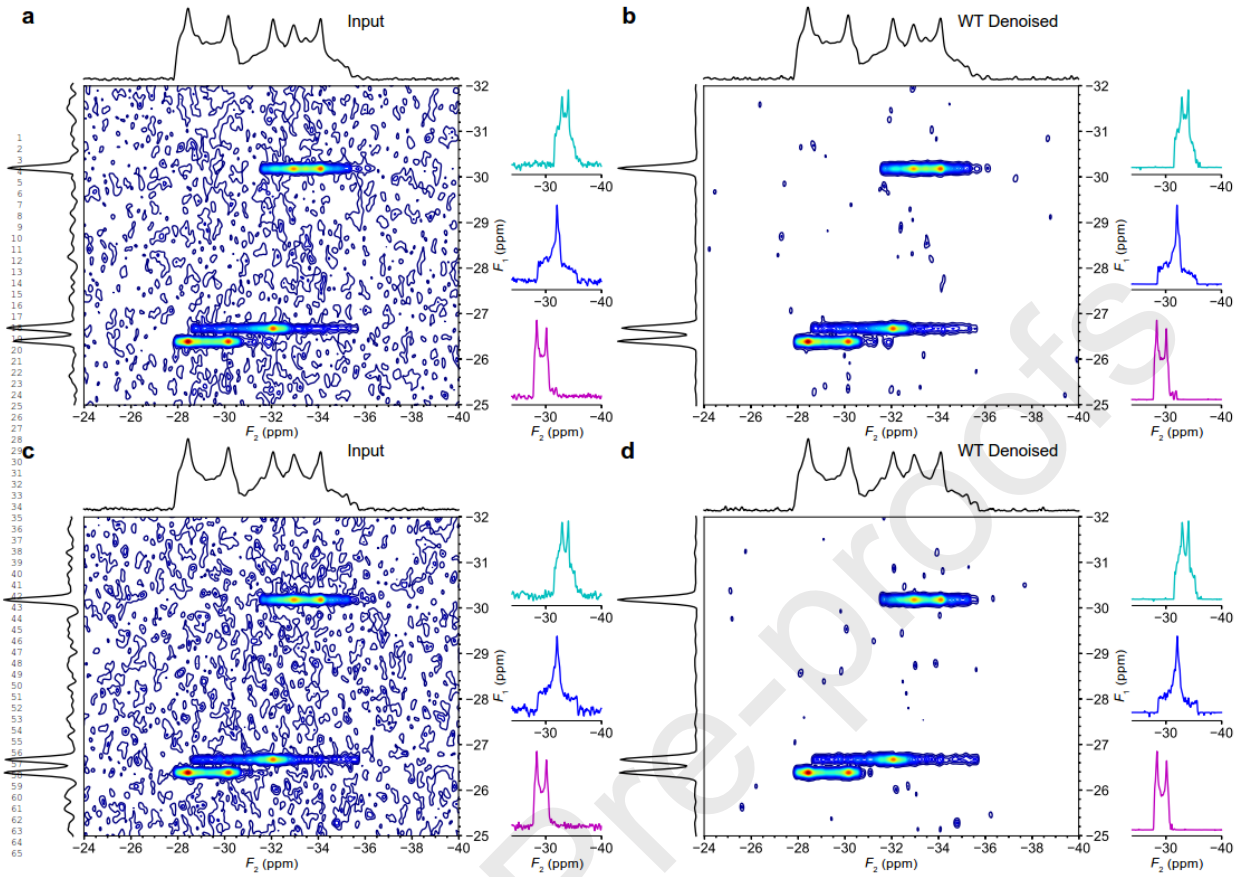


FIG. 9.

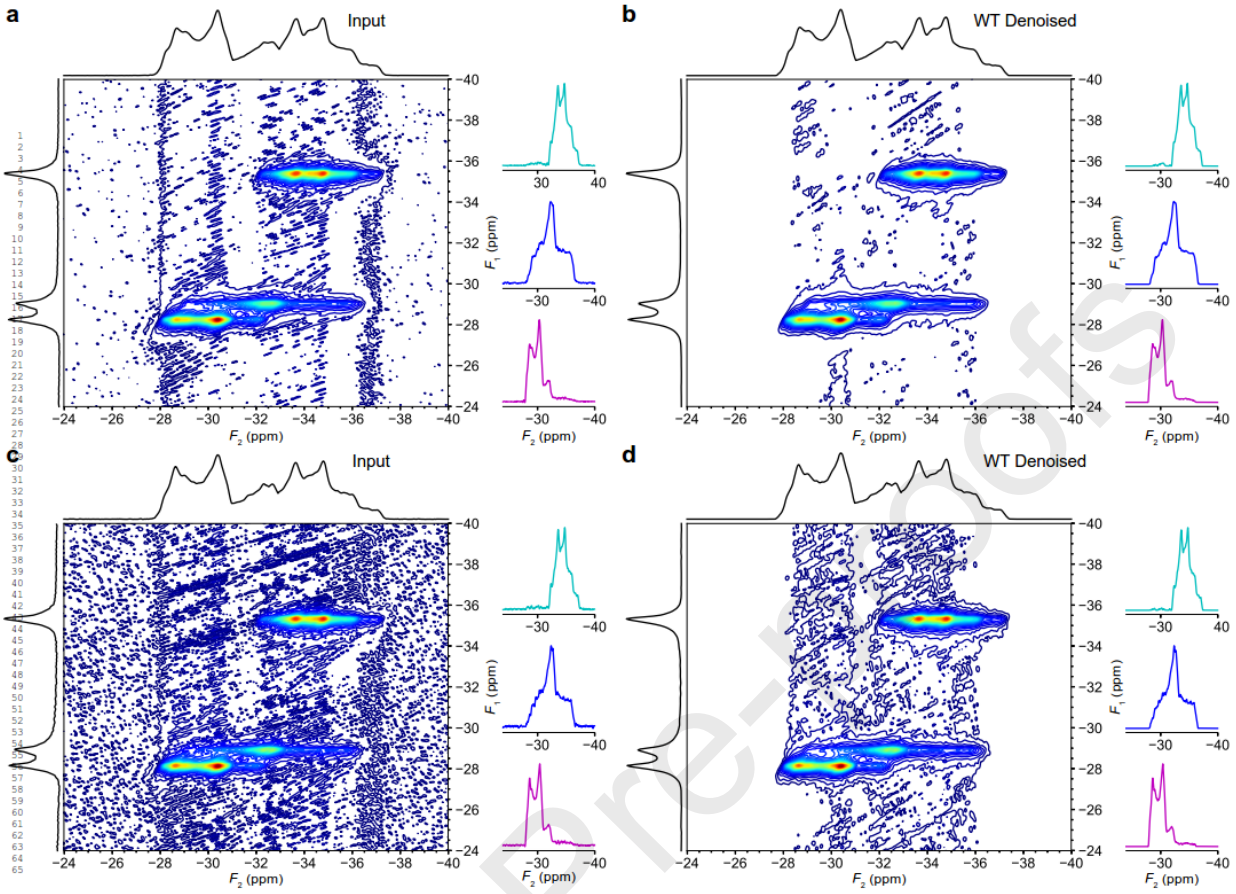


FIG. 10.