Supplementary Material

Manuscript Title: A Robust, Deep Learning-Based Analysis of Time-Domain Signals for NMR Spectroscopy

S1. Model parameter selection and optimization

In our deep learning model, both the input and output vectors contain 2,048 data points, representing time-domain Free Induction Decay (FID) data or frequency-domain NMR spectra. To accurately map FID to spectrum (or peak list), we employ a network architecture with dense layers, each containing 2,048 neurons, matching the input size to ensure sufficient representational power. A shallow configuration of one or two dense layers is used, as deeper architecture did not improve performance in this regression task while increasing complexity and training time.

The Exponential Linear Unit (ELU) activation function is chosen for its advantages in maintaining smooth gradient flow and accelerating convergence. We use the Adam optimizer, known for its adaptive learning rates and efficiency in high-dimensional parameter spaces, with a learning rate of 0.001 providing reliable convergence. Mean Squared Error (MSE) serves as the loss function, aligning directly with the regression objective of minimizing spectral differences. Baseline architecture and hyperparameters were selected based on standard practices for regression tasks, including ELU activations, Adam optimizer, and MSE loss. Subsequent iterative refinements involved grid searches and sensitivity analyses. For example, we tested various learning rates (0.001, 0.0005, 0.0001) and adjusted early stopping patience values (3–10 epochs), with a patience of 5 epochs providing a good balance between training efficiency and convergence stability. Experiments with kernel initializers confirmed He normal as optimal for variance control, while dense layer widths of 2,048 neurons consistently delivered robust performance.

S2. Processing time

For the given input and output vectors, training requires approximately 1 second per epoch, comprising 250 steps (batch size of 32, with a dataset of 8,000–10,000 samples). A typical run of 100 epochs would take around 100 seconds (approximately 1.5 minutes). These timings were obtained on a Mac mini with an Apple M2 Pro chipset and 16 GB of integrated RAM.

Validation after each epoch completes within 1–2 ms per step, while evaluation and predictions are typically processed in 0–2 ms per batch. With early stopping and optimized configurations, the total training time, including validation and evaluation, would be about 2–3 minutes for 20 epochs on standard hardware. Once training and validation are complete, the model can analyze new input data and produce results in just 1–2 ms.

S3. Difference between the deep learning predictions and the traditional FT-derived spectra

We evaluated the difference between the deep learning predictions and the traditional FT-derived spectra using standard error metrics. The Root Mean Squared Error (RMSE) of approximately 0.0305 indicates that, on average, the predicted intensities deviate from the FT reference by about 3% of the normalized scale. Similarly, the Mean Absolute Error (MAE) of about 0.0178 (less than 2% deviation) confirms that the model's predictions closely align with

the conventional FT results. These quantitative measures provide a clear, numerical illustration of the prediction accuracy and demonstrate that the network's outputs are highly consistent with the established FT method.

S4. Structure of the neural networks

Figure S1 shows the structure of neural networks in this study. The network consists of an input layer that takes Free Induction Decay (FID) signals, followed by multiple hidden layers that process the input data, and an output layer that generates the predicted NMR spectrum and peak list.

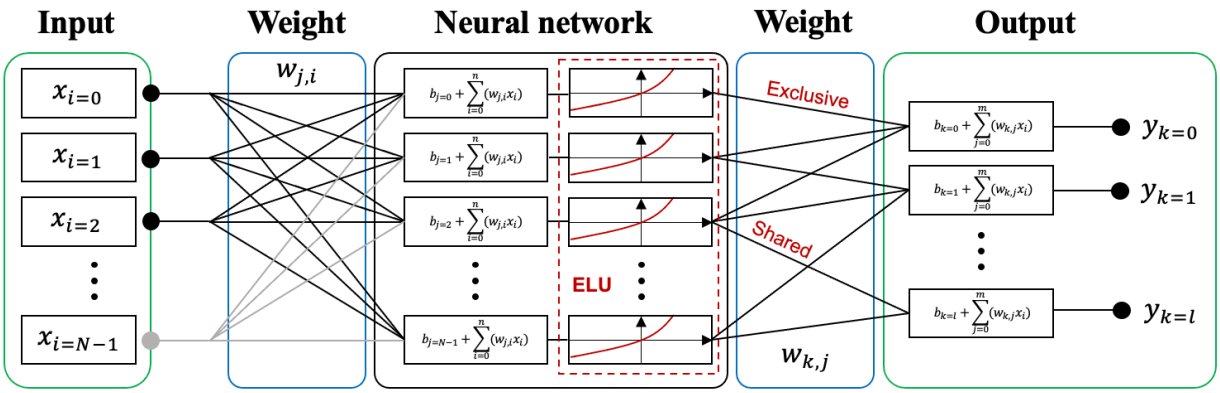

Figure S1. Schematic diagram of the neural network architecture, showing the layer structure used in the model.

S5. Group delay

Group delay arises from digital filtering, especially in modern spectrometers like Bruker. These filters reduce noise but can introduce a phase shift, causing varying delays for different frequency components of the FID signal. This may distort the time-domain signal, affecting the accuracy of peak positions and intensities in the resulting spectrum. Figure S2 shows the generation of group delay in our FID datasets.

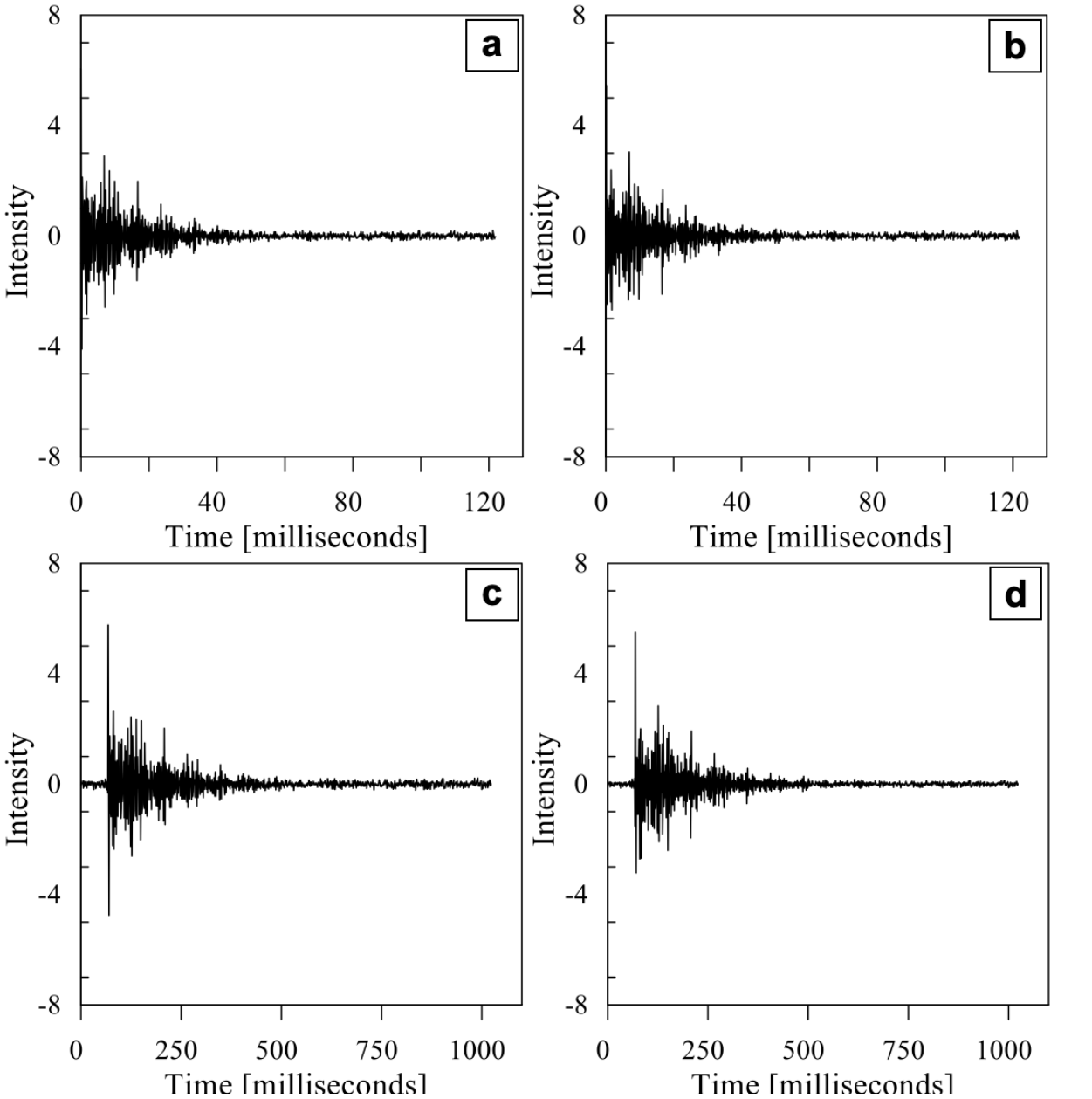

Figure S2. Free Induction Decay (FID) signals. (a) Real part of the FID (cosine component) and (b) imaginary part of the FID (sine component). (c) Real (cosine) and (d) imaginary (sine) parts of the FID after applying group delay, showing how the delay alters the phase and timing of the signal components.

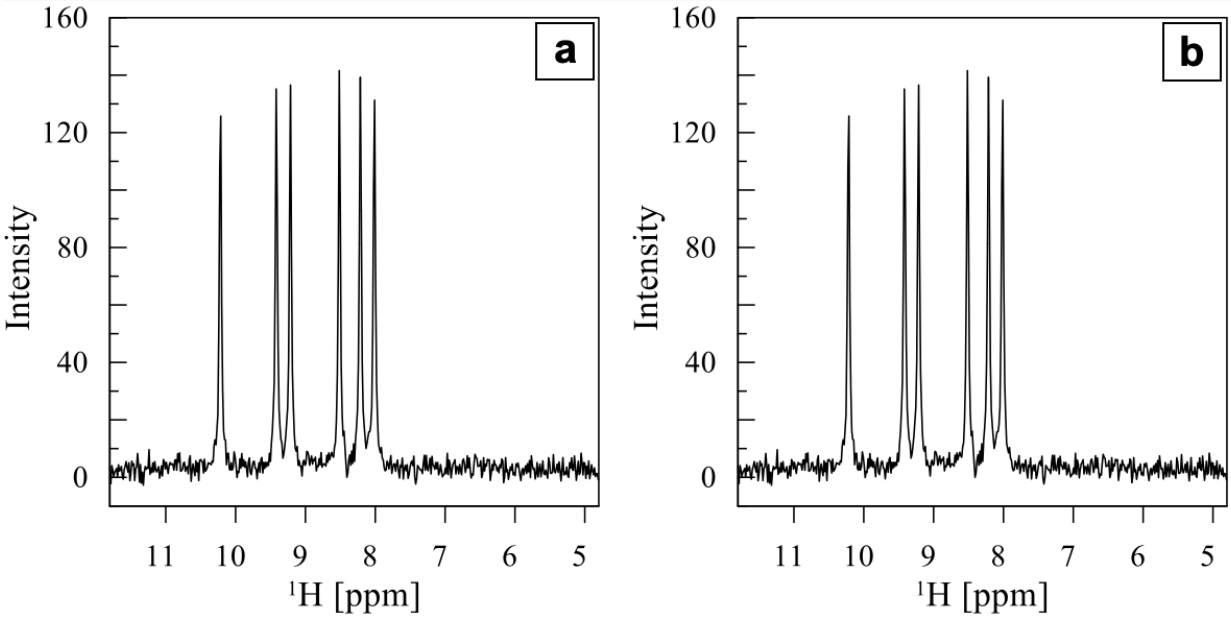

Figure S3. Comparison of spectra derived from (a) the original FID signal and (b) the group delayed FID.

To enable the neural network to predict spectra and peak lists from both FIDs with and without group delay, we trained it using both datasets (Figure S3). This ensures that the model can handle variations in signals due to group delay.

S6. Selection of thresholds

We employed two thresholds in our workflow, 0.2 to eliminate minor baseline fluctuations and 1.5 to isolate exceptionally strong signals. During pilot testing with synthetic FIDs, thresholds set at or below 1.2 resulted in moderate noise or minor peaks being misclassified as "strong," leading to an increased occurrence of false positives. The model frequently misidentified medium-intensity artifacts as legitimate signals. Increasing the threshold to 1.5 effectively removed these extraneous peaks, significantly reducing false positives and enhancing the clarity of prominent signals. The lower threshold of 0.2 was sufficient to filter baseline variations, while the higher threshold of 1.5 targeted only the most intense peaks. Testing this approach on a dataset of 10,000 FIDs demonstrated that a threshold of 1.5 provided an optimal balance. Thresholds below 1.5 (e.g., 1.0 or 1.2) resulted in the misclassification of moderate peaks as "strong," thereby obscuring the distinction between genuine resonances and background noise. Conversely, thresholds above 1.5 (e.g., 1.8 or 2.0) excluded legitimate, moderately strong peaks, leading to an underrepresentation of valid signals.

In the synthetic spectra, most high-intensity signals, particularly those closely resembling experimental resonances, exceeded 1.5 in normalized units. Consequently, the threshold of 1.5 aligns well with the intrinsic amplitude distribution of the dataset, providing a clear separation between strong signals and background or moderate-intensity features.

S7. Generation of noise-free spectra

To improve the predictability of the neural networks, we generated both traditional FT-based spectra and noise-free post-processed spectra for training. Figure S4 shows the differences between these spectra.

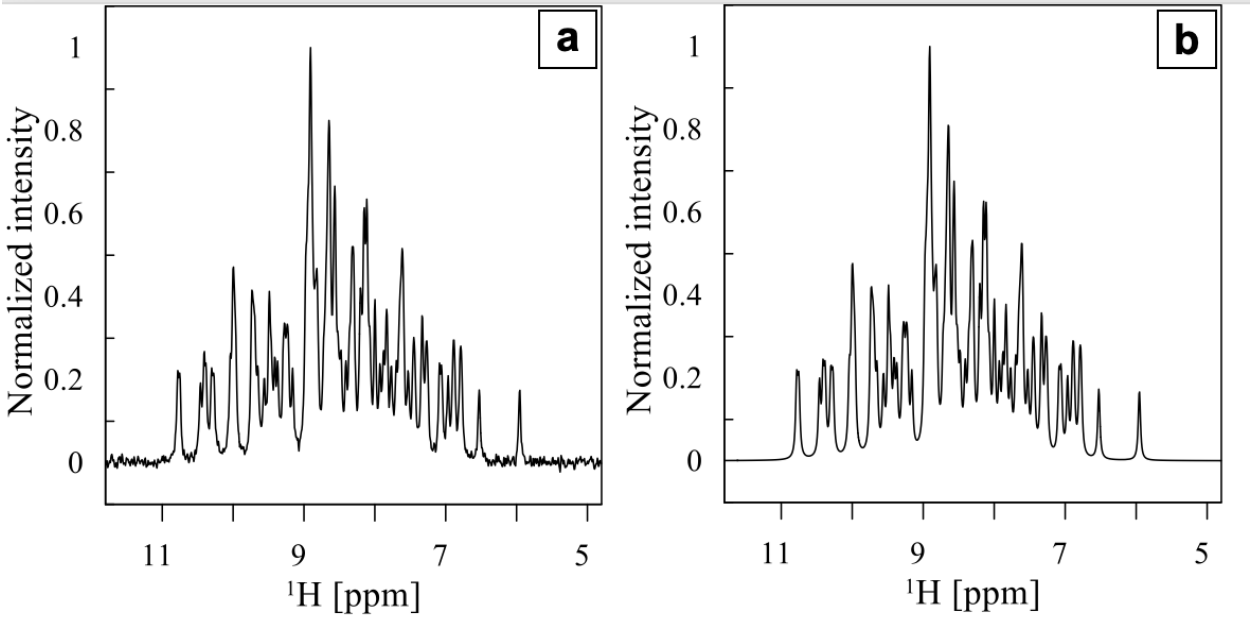

Figure S4. (a) Spectrum processed using standard techniques from the Free Induction Decay (FID) signal, and (b) its corresponding noise-free spectrum.

S8. Model testing with complex overlap spectra

To evaluate the presented methodology under realistic conditions, we designed an additional experiment using an enhanced model. Unlike the original model, which generated random FIDs with fixed or limited peak overlap and a single noise level, the new model simulates complex spectral overlap. This model introduces noise levels randomly selected between 10% and 50% of the signal amplitude, testing the network's robustness in handling low-sensitivity scenarios, such as high noise or low-concentration samples.

The simulation involved generating up to 200 resonances randomly to emulate complex peak overlap, with noise levels ranging from 10% to 50% of the mean signal intensity to replicate low-sensitivity conditions.

Figures 5 (main text) and S5 present the results of the additional experiment, demonstrating strong agreement between the two spectra and the absence of any systematic bias in the predictions.

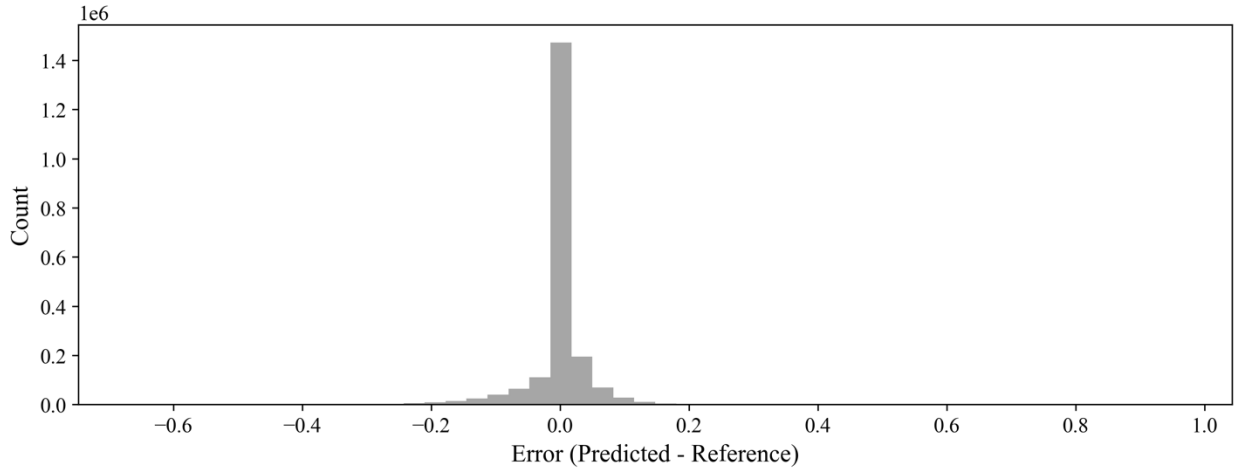

Figure S5. The distribution of errors on the test set, generated under conditions of peak overlap and random noise, divided into 50 bins.

S9. Reduction of relaxation effects

Relaxation including T1 (i.e., Spin-Lattice) and T2 (i.e., Spin-Spin) processes, causes the phase coherence of spins to diminish over time, leading to the gradual decay of the NMR signal. This loss of signal amplitude is observed in the FID. By utilizing neural networks to analyze FID data, it becomes possible to isolate and reconstruct the clean frequency components of the signal, potentially improving the accuracy and resolution of the resulting spectra compared to traditional FT-based methods.

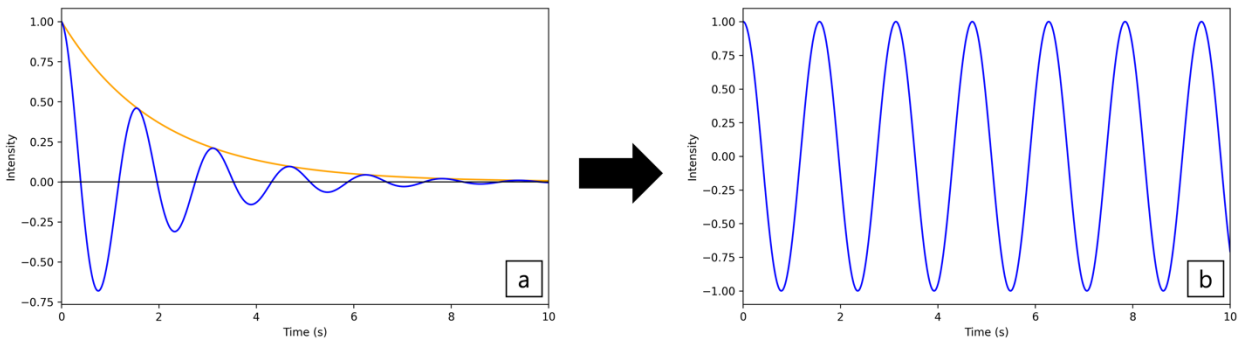

Figure S6. (a) Signal decay due to transverse relaxation (i.e., damped oscillation), with the orange curve representing the exponential decay and the blue curve showing the oscillatory signal with decaying amplitude. (b) The same oscillatory signal without decay showing the absence of relaxation, where the amplitude remains constant over time.