Supplementary information to: Resolution Enhancement in NMR Spectra by Deconvolution with Compressed Sensing Reconstruction

Krzysztof Kazimierczuk[1], Paweł Kasprzak[1,2], Panagiota S. Georgoulia[3], Irena Matečko-Burmann[4,5], Björn M. Burmann[3,5], Linnéa Isaksson[3], Emil Gustavsson[6], Sebastian Westenhoff[3], and Vladislav Yu. Orekhov[* 3,6]

[1] *Centre of New Technologies, University of Warsaw, Banacha 2C, 02-097 Warsaw, Poland*
[2] *Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warsaw, Poland.*
[3] *Department of Chemistry and Molecular Biology, University of Gothenburg, 405 30 Gothenburg, Sweden*
[4] *Department of Psychiatry and Neurochemistry, University of Gothenburg, Gothenburg 405 30, Sweden*
[5] *Wallenberg Centre for Molecular and Translational Medicine, University of Gothenburg, Gothenburg 405 30, Sweden*
[6] *Swedish NMR center, University of Gothenburg, Box 465, 405 30, Gothenburg, Sweden, E-mail: vladislav.orekhov@nmr.gu.se*

1 Deconvolution with Iteratively Re-weighted Least Squares

Let f be the the subsampled NMR (complex) signal measured at $\{t_1, \ldots, t_k\}$ and viewed as the column vector $f \in \mathbb{C}^k$. The sampling schedule $\{t_1, \ldots, t_k\}$ is a fixed subset of the full sampling grid $\{l\Delta t : l = 0, 1, \ldots, n-1\}$. The CS methodology when applied in NMR context provides the algorithms and theory which enable for the recovery of the NMR spectrum $s \in \mathbb{C}^n$ from the sub-sampled signal, even if $k \ll n$. The CS problem is formulated in terms of the measurement matrix $F \in M_{k\times n}(\mathbb{C})$. In the NMR application of CS the matrix F consists of rows of the $n \times n$ inverse Fourier matrix corresponding to the sampling schedule and we have

$$f = Fs. \tag{1}$$

Note that Eq. (1) with known f and unknown s and $k < n$, has infinitely many solutions. The fundamental insight of the CS theory [1] specifies the NMR spectrum s as the unique solution of the convex optimization problem

$$s = \underset{x\in\mathbb{C}^n}{\arg\min} \left(\|Fx - f\|_2^2 + \lambda \|x\|_1 \right). \tag{2}$$

The first term in this sum promotes the consistency of tested x with the measured data whereas, the second term promotes the sparseness of x and $\lambda > 0$ fixes the balance between the two.

The Eq. (2) can be modified to take better account of the measurement noise. We can consider two cases:

- Unstructured noise: the covariance noise matrix Σ is diagonal and isotropic, $\Sigma = \sigma^2 \mathbb{1}$ where σ is the standard deviation of the noise. In order to take into an account the level of noise in the reconstruction framework (2) one sets λ proportional to σ^2.

- Structured noise: the covariance noise matrix Σ is not necessarily diagonal. The ℓ_2 norm used in the data consistency term should be replaced by its weighted version $\|Fx - f\|_Q^2$, where $Q = \Sigma^{-1}$ and for a given complex positive definite matrix G and a complex vector y we define $\|y\|_G^2 = y^\dagger G y$ where $y^\dagger$ is the conjugated transpose of y. In order to justify the proposed modification of the data consistency term let us discuss an instructive example. For that matter consider two independent measurements $f_1 = f(t_1), f_2 = f(t_2)$ of the signal in which the noise level of $f(t_1)$ is two times smaller than that of $f(t_2)$. The corresponding covariance matrix of Σ is of the form

$$\Sigma = \begin{bmatrix} \sigma^2 & 0 \\ 0 & 4\sigma^2 \end{bmatrix}$$

where σ is the noise level entering $f(t_1)$. Let us denote $f^\# = Fx$ the signal corresponding to the spectrum x. The consistency term

$$\|Fx - f\|_2^2 = |f_1^\# - f_1|^2 + |f_2^\# - f_2|^2$$

entering the standard formulation of the CS-problem is replaced by

$$\|Fx - f\|_Q^2 = (f^{\#\dagger} - f^\dagger)\Sigma^{-1}(f^\# - f) = \frac{1}{\sigma^2}|f_1^\# - f_1|^2 + \frac{1}{4\sigma^2}|f_2^\# - f_2|^2.$$

Our modification introduces the weights in the data consistency term that correctly reflect the noise level of the corresponding measurements. Points with larger noise enter the sum with smaller weights. The above discussion and conclusion easily generalizes to larger number of independent measurements. In case of the non-zero correlations between the noise of the measurements, Σ must be first diagonalized in an appropriate orthonormal basis and the above justification can be then repeated.

Let us note that for the unstructured noise ($\Sigma = \sigma^2 \mathbb{1}$) we have

$$\|Fx - f\|_Q^2 = \sigma^{-2}\|Fx - f\|_2^2 \tag{3}$$

and thus the cases of structured and unstructured noise are consistent:

$$\begin{aligned} \arg\min_{x\in\mathbb{C}^n} \left(\|Fx - f\|_Q^2 + \lambda\|x\|_1\right) &= \arg\min_{x\in\mathbb{C}^n} \left(\sigma^{-2}\|Fx - f\|_2^2 + \lambda\|x\|_1\right) \\ &= \arg\min_{x\in\mathbb{C}^n} \left(\|Fx - f\|_2^2 + \sigma^2\lambda\|x\|_1\right). \end{aligned}$$

Let us also note that in each case we can absorb λ into Σ by the possible scaling $\Sigma \rightsquigarrow \lambda\Sigma$.

Now, we will describe algorithm known as Iteratively Reweighted Least Squares (IRLS) [2, 3, 4] dedicated for the solution of

$$s = \arg\min_{x\in\mathbb{C}^n} \left(\|Fx - f\|_Q^2 + \|x\|_1\right). \tag{4}$$

Note first, that the ℓ_1 norm $\|x\|_1 = \sum_{i=1}^n |x_i|$ can be well approximated by the weighted ℓ_2 -norm $\sum_{i=1}^n |w_i x_i|^2$, where the weights $w_i = |x_i|^{-1/2}$ are regularized for very small x_i's. This seemingly trivial observation is the starting point of IRLS[5], which is an iterative procedure that solves the quadratic problem

$$s_l = \arg\min_{x\in\mathbb{C}^n} \|Fx - f\|_Q^2 + \|w_l x\|_2^2 \tag{5}$$

where l is the iteration loop number and w_l is the weight vector corresponding to s_{l-1} as described above. Thus defining $D_l = \mathrm{diag}(d_1, \dots d_n)$ where $d_j = \frac{1}{|s_{l-1,j}|+\varepsilon}$ (we write $s_{l-1,j}$ for the jth component of the vector s_{l-1}), Eq.(5) can be written in the form of generalized Tikhonov regularization

$$s_l = \arg\min_{x\in\mathbb{C}^n} \|Fx - f\|_Q^2 + \|x\|_{D_l}^2 \tag{6}$$

and the latter can be solved explicitly

$$s_l = (F^* Q F + D_l)^{-1} F^* Q f.$$

J-coupling in IRLS

The J-modulation of the NMR signal f is represented in vector language by a complex vector $M \in \mathbb{C}^k$. For example M corresponding to the J-coupling considered in the main text (i.e. approximately the same for all components) is of the form $M(t_j) = \cos(\pi J t_j)$. The unmodulated version $\tilde{f}$ of f is defined by the equality $f = M\tilde{f}$, or more precisely $f(t_i) = M(t_i)\tilde{f}(t_i)$ where $i \in \{1, \ldots, k\}$. Let $C \in M_{k\times k}(\mathbb{C})$ denote the modulation (diagonal) matrix

$$C = \mathrm{diag}(M(t_1), \ldots, M(t_k)).$$

The relation between $\tilde{f}$, f and the spectrum $\tilde{s}$ corresponding to unmodulated signal $\tilde{f}$ is of the form

$$f = C\tilde{f} = CF\tilde{s}.$$

Note, that while the noise in the measured signal f is unstructured, the noise of $\tilde{f}$ is structured (since $\tilde{f} = C^{-1}f$).

Denoting the spectrum related to modulated signal f by s we have $f = Fs$. Remarkably, the spectrum $\tilde{s}$ is sparser than s for typical modulations encountered in NMR. For example in the simplest case of one dimensional J-modulated signal, singlets in $\tilde{s}$ are doubled in s, i.e. the number $\tilde{m}$ of significant elements of the spectrum $\tilde{s}$ is approximately half of the number m of the significant component of s. The standard estimation for the number of measurements k required for the exact CS reconstruction is[6]

$$k \sim m \log(n). \tag{7}$$

This estimation and the preceding remark show that that number of measurements $\tilde{k}$ required for the recovery of $\tilde{s}$ gets divided by the factor 2 when compared with k required for the recovery of s. The above observations motivate the development of a CS methodology dedicated to the signal in the presence of modulations, which we formulate in what follows.

Assuming that the J-modulated (measured) signal f is corrupted by the unstructured noise with covariance matrix Σ, the de-modulated signal $\tilde{f} = C^{-1}f$ is corrupted by the structured noise with the covariance matrix $\tilde{\Sigma} = C^{-1}\Sigma C^{\dagger -1}$. In particular the weighting matrix $\tilde{Q} = \tilde{\Sigma}^{-1}$ is equal to $C^\dagger QC$ and we get

$$\begin{aligned}
\|Fx - \tilde{f}\|^2_{\tilde{Q}} &= (Fx - \tilde{f})^\dagger C^\dagger QC(Fx - f) \\
&= (CFx - C\tilde{f})^\dagger Q(CFx - C\tilde{f}) \\
&= (CFx - f)^\dagger Q(CFx - f) \\
&= \|CFx - f\|^2_Q.
\end{aligned}$$

This computation shows that the solution of minimization problem

$$\tilde{s} = \underset{x\in\mathbb{C}^n}{\arg\min} \left(\|Fx - \tilde{f}\|^2_{\tilde{Q}} + \|x\|_1 \right)$$

coincides with that corresponding to

$$\tilde{s} = \underset{x\in\mathbb{C}^n}{\arg\min} \left(\|CFx - f\|^2_Q + \|x\|_1 \right)$$

and the latter can be found by the IRLS in the iterative procedure

$$\tilde{s}_l = \underset{x\in\mathbb{C}^n}{\arg\min} \left(\|CFx - f\|^2_Q + \|x\|^2_{D_l} \right). \tag{8}$$

In the resulting spectrum the multiplets reflecting the modulation (e.g. doublets) will be replaced by singlets. Keeping in mind the fact that in the standard experimental setup the measurement

noise of f is described by $\Sigma = \sigma^2 \mathbb{1}$, and substituting $\lambda = \sigma^2$ as explained above, we observe that Eq.(8) boils down to the iterative procedure based on generalized Tikhonov regularization

$$\tilde{s}_l = \underset{x \in \mathbb{C}^n}{\arg\min} \left(\|CFx - f\|_2^2 + \lambda \|x\|_{D_l}^2 \right). \tag{9}$$

Remarkably this is the IRLS algorithm in the orthodox form (c.f. (2)) applied to

$$\tilde{s} = \underset{x \in \mathbb{C}^n}{\arg\min} \left(\|CFx - f\|_2^2 + \lambda \|x\|_1 \right), \tag{10}$$

which in turn may be viewed as mathematical formulation of the problem of CS-type of finding a sparse spectrum $\tilde{s}$ whose consistency with the measurement vector f is given by the measurement matrix CF: $f = CF\tilde{s}$.

Summarizing, in the above considerations we explained the following two aspects:

- we formulate the CS-problem suitable for the case of noisy measurements described by the structured measurements noise; we also describe the solution of this problem and relate it with the Tikhonov regularization;
- we applied the modified CS-methodology to the J-modulated signals in NMR and explain the advantages of our approach when compared to the orthodox version of CS.

2 Protein samples and NMR spectroscopy

A [U-^{2}H,^{15}N,^{13}C] labeled sample of the monomeric photosensory module (57 kDa) *Dr*BphP$_{PSM}$ from *Deinococcus radiodurans* was produced exactly as described in our previous study[7, 8]. The monomeric variant contains three mutations, which disrupt the dimer interface: F145S, L311E, and L314E[9].

A 3D BEST-TROSY-HNCA experiment [10] was recorded for the *Dr*BphP$_{PSM}$ sample during 38.5 hours with 2930 mildly relaxation-matched NUS points (probability density exponentially biased with T_2 = 70 ms) in the $^{13}C^{\alpha}$ dimension. The spectral widths (acquisition times) were 12.8 kHz (80 ms), 2.9 kHz (22 ms), and 6.0 kHz (42.4 ms), for ^{1}H, ^{15}N, and ^{13}C spectral dimensions, respectively. For the processing, the original NUS data set was sub-sampled to create the following data sets:

- The "traditional" low resolution spectrum processed using IRLS algorithm without the virtual decoupling, 1170 NUS points (15.4 hours of measurement time) were retained from the original NUS data with maximal evolution time of 14 ms for the $^{13}C^{\alpha}$ dimension (Figure S1a)).

- The high resolution spectra processed using IRLS (undecoupled) and the region-selective D-IRLS algorithm (decoupled). The original NUS data was sub-sampled down to 1200 NUS points (15.8 hours of measurement time). The resulting sampling probability distribution for this spectrum corresponded to the sampling in the ^{13}C dimension matched to the J-coupling and mildly biased for the relaxation (T_2 = 70 ms, J=35 Hz), i.e. proportional to $\exp(-t/T_2)|\cos(\pi Jt)|$, with additional elimination of the points whose values of $|\cos(\pi Jt)|$ were less than 0.2 (Figure S1b)).

A [U-^{2}H,^{15}N,^{13}C] labeled sample of the longest human Tau protein isoform hTau40 with 441 residues was prepared as following: full length hTau40 with an amino-terminal His$_6$-SUMO-Tag (in a modified pET28b plasmid, Genescript) was expressed in *E. coli* BL21(λDE3) Star™(Novagen) cells. [U-^{2}H,^{15}N,^{13}C] isotope (Merck) enriched protein was produced using 2xM9 minimal medium[11] supplemented with $^{15}NH_4Cl$ and D-(^{2}H/^{13}C)- glucose as the sole nitrogen and carbon sources, respectively, in D_2O. The cells were grown at 37 °C until an $OD_{600} \approx 0.8$. Expression was induced by addition of 1mM isopropyl-thiogalactoside (IPTG) (Thermo Scientific) for 16h at 22 °C. Cells were harvested by centrifugation and subsequently resuspended in lysis buffer (20mM NaPi, 500mM NaCl, pH 7.8), and lysed by an Emulsiflex C3 (Avestin) homogenizer. Cleared lysate was purified with HisTrap HP column (GE Healthcare). Fractions containing hTau40 were pooled and dialyzed over-night, against human SenP1 cleavage buffer (20 mM TrisHCl, 150 mM NaCl, 1 mM DTT, pH 7.8). After dialysis SenP1 protease (Addgene #16356)[12] was added and the enzymatic cleavage was performed for 4 hours at room temperature, followed by a second HisTrap HP column step. Fractions containing cleaved hTau40 in the flow-through were collected, concentrated, and subsequently purified by gel filtration using a HiLoad 10/60 200 pg (GE Healthcare) pre-equilibrated with NMR buffer (25mM NaPi, 50mM NaCl, 1mM EDTA, pH 6.9). The pure hTau40 fractions were concentrated to about 500 μM, flash frozen in liquid nitrogen, and stored at -80 °C till usage.

A 3D BEST-TROSY-HN$^{13}C^{\alpha}$ experiment [10] for hTau40 was recorded and processed similar to the *Dr*BphP$_{PSM}$ spectrum described above. Namely, it was acquired during 13 hours with 1552 NUS points using spectral width (acquisition times) of 9.6 kHz (106 ms), 2.9 kHz (22 ms), and 6.0 kHz (42.4 ms), for ^{1}H, ^{15}N, and ^{13}C spectral dimensions, respectively. For the processing, the originally random NUS data set was sub-sampled to create the following data sets:

- The "traditional" low resolution spectrum was processed using IRLS algorithm without the virtual decoupling, 555 NUS points (4.6 hours of measurement time) were retained from the original NUS data with a maximal evolution time of 14 ms for the $^{13}C^{\alpha}$ dimension (Figure S1c)).

- The high resolution spectra processed using IRLS (undecoupled) and the region-selective D-IRLS algorithm (decoupled). The original NUS data was sub-sampled down to 1200 NUS

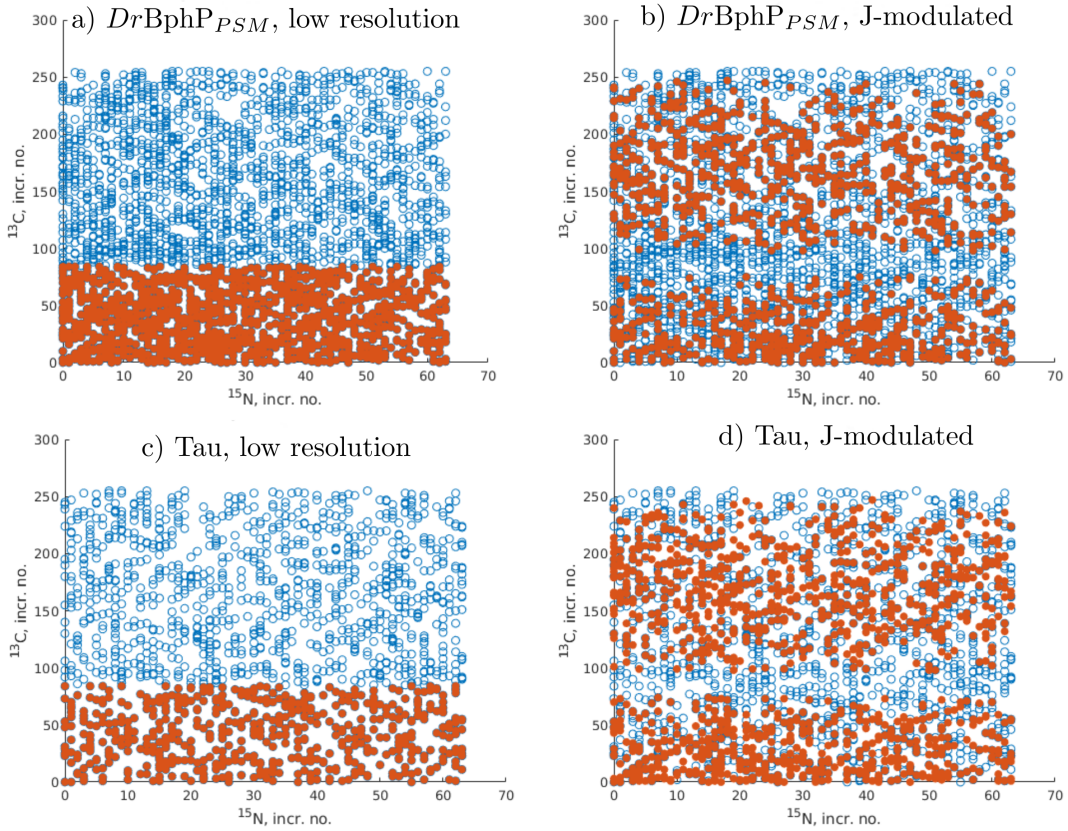

Figure S1: Sampling schedules used in the original experiments (blue) and sub-sampled data (red). The original schedules include 2930 points for DrBphP$_{PSM}$ and 1552 for Tau and follow the relaxation-matched NUS scheme with T_2 = 200 ms for Tau c) and d) and T_2 = 70 ms for DrBphP$_{PSM}$ c) and d). Schedules used for calculation of the "low-resolution" spectra, shown in panels a) and c) for DrBphP$_{PSM}$ and Tau, respectively, were created by truncating the ^{13}C dimension to 14 ms which resulted in 1170 a) and 555 c) points. The "J-modulated" schedules shown in panels b) and d), for DrBphP$_{PSM}$ and Tau, respectively, were created by selecting 1200 points that matched the $|\cos(\pi Jt)|$ envelope in the ^{13}C dimension (with elimination of points for which $|\cos(\pi Jt)| < 0.2$).

points (10 hours of measurement time). The resulting sampling probability distribution for this spectrum corresponded to the sampling in the ^{13}C dimension matched to the J-coupling and mildly biased for the relaxation (T_2 = 200 ms, J=35 Hz), i.e. proportional to $\exp(-t/T_2)\cos(\pi Jt)$, with additional elimination of the points whose values of $|\cos(\pi Jt)|$ were less than 0.2 (Figure S1d)).

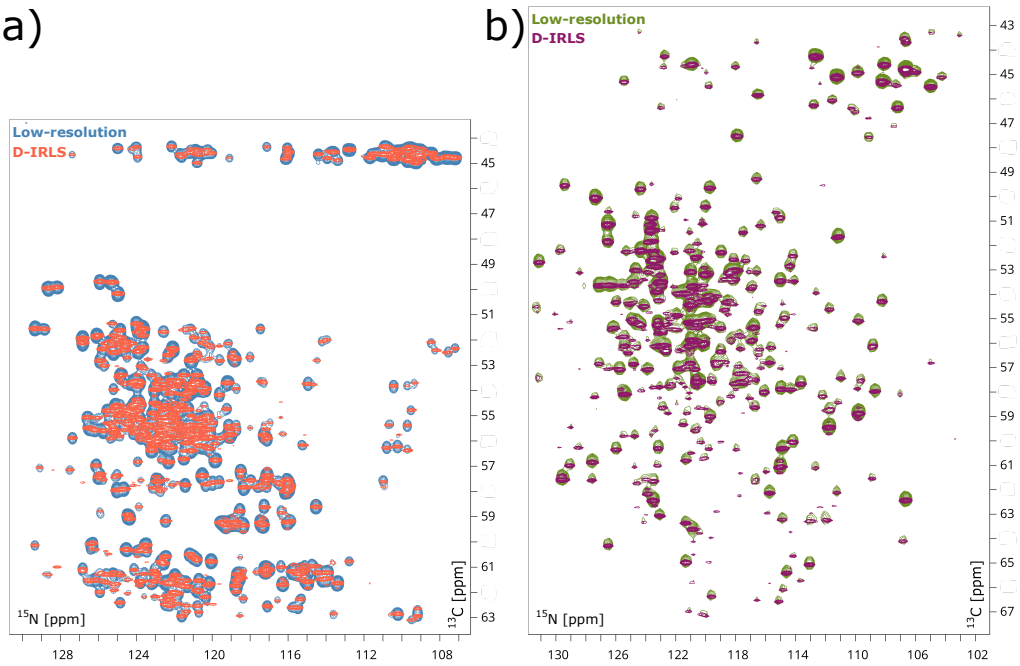

Figure S2: ^{15}N/^{13}C projections of 3D HNCA spectra of a) Tau and b) *Dr*BphP$_{PSM}$ proteins. The plots show superimposed traditional "low-resolution" and high-resolution deconvoluted spectra with 14 ms and 42.4 ms of maximum evolution time in the ^{13}C dimension, respectively.

3 Simulations with synthetic peaks

Synthetic peaks were injected into the spectrum of Tau protein in order to estimate accuracy of the spectra reconstructions. Comparison was performed between the reconstruction obtained using two types of NUS sampling schedules:

- matched to both $^{13}C^{\alpha}$ transverse relaxation and J-modulation
- matched to the $^{13}C^{\alpha}$ relaxation only

and three calculations modes:

- high resolution spectrum with $^{13}C^{\alpha}$ maximum evolution time of 42 ms - IRLS reconstruction without the deconvolution
- high resolution spectrum with $^{13}C^{\alpha}$ maximum evolution time of 42 ms - IRLS reconstruction with the deconvolution (D-IRLS) for the whole spectrum
- region-selective processing with IRLS for the Gly region (<45 ppm in $^{13}C^{\alpha}$) and D-IRLS for the rest of the spectrum (as in Figure 1 in the main text).

Each version of the reconstruction was calculated 15 times using selected with a corresponding random distribution (sub-sampled) NUS data sets from the larger pool of measured data (total 1552 flat-random NUS). In each calculation 20 peaks, including 5 Gly-type peaks, were injected with random positions (without overlap between each other and the existing Tau protein peaks) and intensities varying in the range 0.05-1.0 of a typical medium strong peak in the original Tau spectrum. The non-Gly signals were injected as the doublets with random J-coupling value in the range 35±5 Hz. The resulting peaks were automatically picked in the reconstructed spectra using function pkFindROI in nmrPipe software[13]. For the undecoupled spectrum, intensities and frequencies of the picked doublet components were added and averaged, respectively. Thus,

in each variant of the spectrum reconstructions, up to 225 non-Gly synthetic peaks were picked and quantified. Statistics on the peaks in accuracy of the peak intensities and positions is shown in Figure S3, where the comparisons are given for four calculations using 250, 400, 700, and 1000 NUS points.

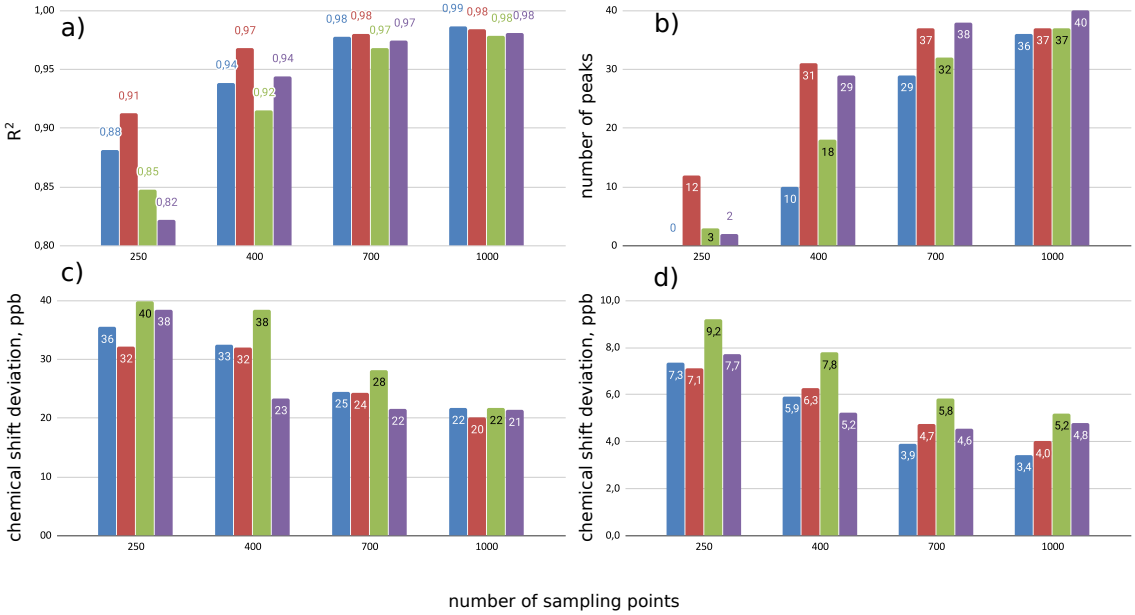

Figure S3: The results of simulations with synthetic peaks injected into the experimental spectrum of Tau protein. The columns correspond to: undecoupled IRLS (blue), D-IRLS with (red) and without (green) Gly-region extraction, all with J-modulated sampling scheme, and D-IRLS with Gly-region extraction with non-modulated sampling (purple). The plots show: a) R^2 coefficients of peak intensity reconstruction, b) number of detected weak peaks out of 42 lowest intensity injected peaks, c) deviation of peak position in ^{15}N dimension, d) deviation of peak position in $^{13}C^{\alpha}$ dimension. Numbers at the tops of the columns indicate their exact heights.

4 Important parameters

Both main text and SI mentioned several important parameters of the deconvolution approach. For clarity, we list them again below.

- **Sampling level.** For the two complex systems used in this work, the 7.3% sampling was sufficient for the IRLS-D of high-resolution HNCA. Even lower levels may be fine for smaller proteins and if spectrum resolution is increased, e.g. for ^{15}N dimension. In practice, however, the signal-to-noise ratio also plays an important role, and the total measurement time should be selected accordingly.

- **The sampling scheme: modulation and cut-off.** The signal-to-noise ratio can be improved by using J-modulated density of sampling points. The modulation also improves robustness of the CS reconstruction by avoiding zero-crossings of the signal. The latter is additionally assured by the hard threshold eliminating time points where the cosine function is lower than the threshold value. In this paper, we used the cosine modulation of 35 Hz and a threshold of 0.2. The setting should work nicely for all protein systems.

- **Deconvolution constant.** Although the J-couplings in the signal may vary up to ±5 Hz, using the average value 35 Hz is fine as long as the line width in the spectrum exceeds the variation range.

- **Non-deconvolved region.** As described in the main text, the deconvolution can be performed selectively in different regions. In the case of HNCA spectrum, the border of the deconvolution region should be set to exclude glycine resonances i.e. at about 45 ppm. As long as no signal is crossed by the border and no singlets are found in the deconvolution region, the parameter specific value does not affect the reconstruction quality.

- **Regularization parameters (λ and ϵ).** The regularization parameters of the IRLS algorithm play a similar role in controlling the sparsity of the spectrum. In the IRLS version implemented in mddnmr software[14] used in this study, ϵ is set automatically. It is also recommended to keep the λ parameter (CS_lambda in the program) to its default value of 1.0. The value may require increase by orders of magnitude if non-modulated NUS schedule was used and, consequently, the time points near the zero-crossings of the J-modulating cosine function used in the deconvloution are extensively sampled.

References

[1] Simon Foucart and Holger Rauhut. *A Mathematical Introduction to Compressive Sensing*. Wiley, 2010.

[2] Rick Chartrand and Wotao Yin. Iteratively reweighted algorithms for compressive sensing. In *ICASSP*, pages 3869–3872. IEEE, 2008.

[3] Krzysztof Kazimierczuk and Vladislav Yu. Orekhov. Accelerated NMR spectroscopy by using compressed sensing. *Angewandte Chemie - International Edition*, 50(24):5556–5559, 2011.

[4] Krzysztof Kazimierczuk and Vladislav Yu. Orekhov. The comparison of convex and non-convex compressed sensing applied in multidimensional NMR. *Journal of Magnetic Resonance*, 223(0):1–10, 2012.

[5] Emmanuel J. Candès, Michael B. Wakin, and Stephen P. Boyd. Enhancing sparsity by reweighted L1 minimization. *Journal of Fourier Analysis and Applications*, 14(5-6):877–905, 2008.

[6] E. J. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, Feb 2006.

[7] Emil Gustavsson, Linnéa Isaksson, Cecilia Persson, Maxim Mayzel, Ulrika Brath, Lidija Vrhovac, Janne A. Ihalainen, B. Göran Karlsson, Vladislav Orekhov, and Sebastian Westenhoff. Modulation of Structural Heterogeneity Controls Phytochrome Photoswitching. *Biophysical Journal*, 118(2):415–421, 2020.

[8] bmrb entry 10.13018/bmr27783.

[9] H. Takala, A. Björling, M. Linna, S. Westenhoff, and J.A. Ihalainen. Light-induced Changes in the Dimerization Interface of Bacteriophytochrome. *The Journal of Biological Chemistry*, 290(26):16383–16392, 2015.

[10] Zsofia Solyom, Melanie Schwarten, Leonhard Geist, Robert Konrat, Dieter Willbold, and Bernhard Brutscher. BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins. *Journal of Biomolecular NMR*, 55(4):311–321, 2013.

[11] Stephan B Azatian, Navneet Kaur, and Michael P Latham. Increasing the buffering capacity of minimal media leads to higher protein yield. *Journal of Biomolecular NMR*, 73(1-2):11–17, 2019.

[12] Jowita Mikolajczyk, Marcin Drag, Miklós Békés, John T Cao, Ze'ev Ronai, and Guy S Salvesen. Small ubiquitin-related modifier (sumo)-specific proteases profiling the specificities and activities of human senps. *Journal of Biological Chemistry*, 282(36):26217–26224, 2007.

[13] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *Journal of Biomolecular NMR*, 6(3):277–293, 1995.

[14] V Yu. Orekhov, V Jaravine, M Mayzel, and K Kazimierczuk. MddNMR - Reconstruction of NMR spectra from NUS signal using MDD and CS, 2004-2020.