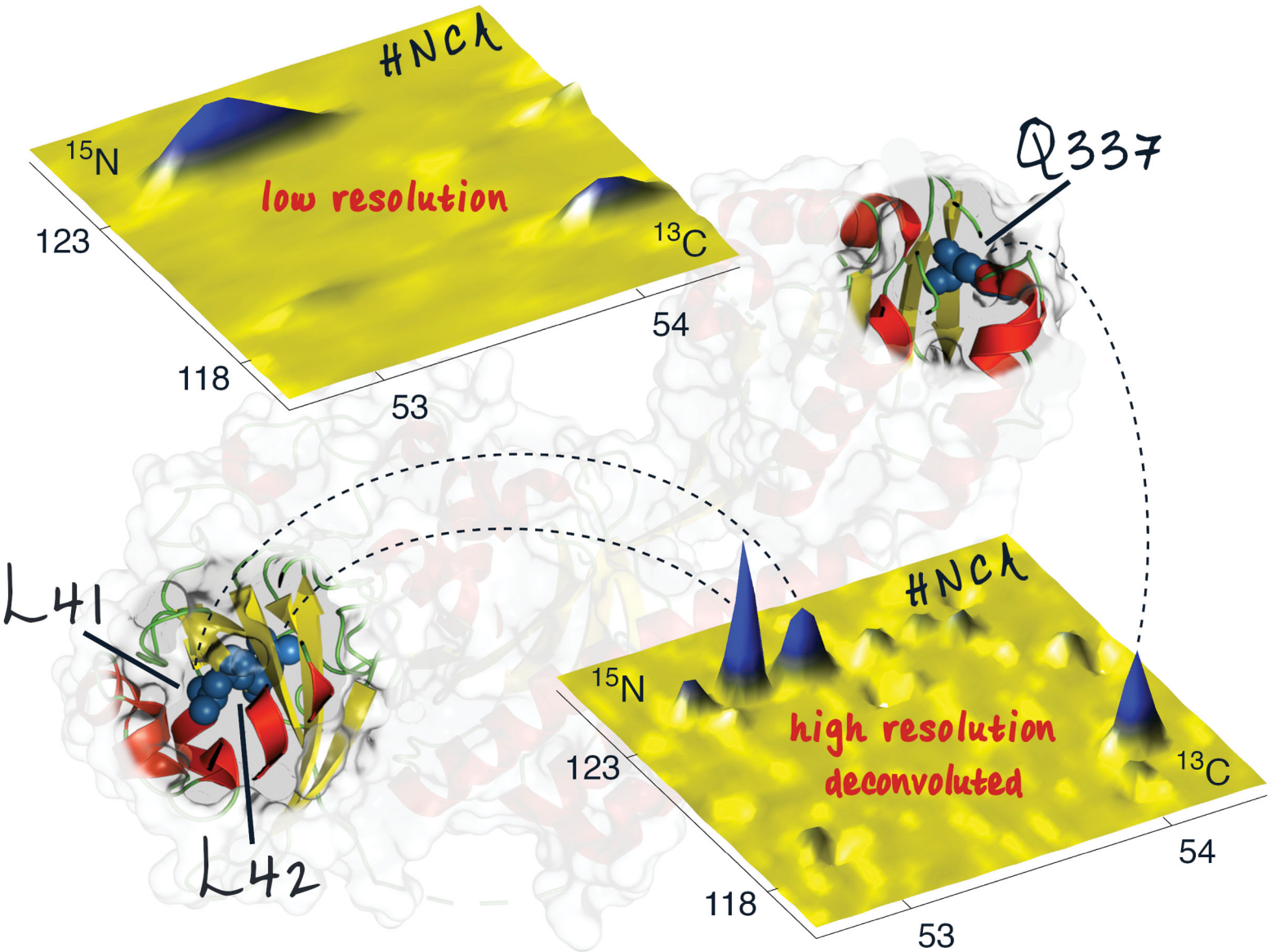
Volume 56
Number 93
4 December 2020
Pages 14505–14730

ChemComm

Chemical Communications

rsc.li/chemcomm



ISSN 1359-7345

ROYAL SOCIETY OF CHEMISTRY

COMMUNICATION

Vladislav Yu. Orekhov *et al.*

Resolution enhancement in NMR spectra by deconvolution with compressed sensing reconstruction

ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Chem. Commun.*, 2020, **56**, 14585

Received 13th September 2020,
Accepted 28th October 2020

DOI: 10.1039/d0cc06188c

rsc.li/chemcomm

Resolution enhancement in NMR spectra by deconvolution with compressed sensing reconstruction†

Krzysztof Kazimierczuk, [a] Paweł Kasprzak, [ab] Panagiota S. Georgoulia, [c] Irena Matečko-Burmann, [de] Björn M. Burmann, [ce] Linnéa Isaksson, [c] Emil Gustavsson, [f] Sebastian Westenhoff [c] and Vladislav Yu. Orekhov *[cf]

NMR spectroscopy is one of the basic tools for molecular structure elucidation. Unfortunately, the resolution of the spectra is often limited by inter-nuclear couplings. The existing workarounds often alleviate the problem by trading it for another deficiency, such as spectral artefacts or difficult sample preparation and, thus, are rarely used. We suggest an approach using the coupling deconvolution in the framework of compressed sensing (CS) spectra processing that leads to a major increase in resolution, sensitivity, and overall quality of NUS reconstruction. A new mathematical description of the decoupling by deconvolution explains the effects of thermal noise and reveals a relation with the underlying assumption of the CS. The gain in resolution and sensitivity for challenging molecular systems is demonstrated for the key HNCA experiment used for protein backbone assignment applied to two large proteins: intrinsically disordered 441-residue Tau and a 509-residue globular bacteriophytochrome fragment. The approach will be valuable in a multitude of chemistry applications, where NMR experiments are compromised by the homonuclear scalar coupling.

Nuclear magnetic resonance (NMR) is among the main analytical techniques allowing atomic-level studies of proteins. The prerequisite step for most protein NMR work is a resonance-specific spectral assignment, *i.e.* association of resonance frequencies with atoms in the protein amino acid chain.[1] The HNCA[2–4] is by far the most sensitive and thus often the only feasible triple resonance experiment that provides sequential connectivities between neighbouring protein residues. In principle, a sequence-specific resonance assignment could be obtained using the HNCA experiment alone. Unfortunately, low signal resolution relative to the dispersion of the protein $^{13}C^{\alpha}$ resonances results in massive ambiguity of the assignment even for relatively small proteins. For large systems with many amino acid residues as well as for intrinsically disordered proteins (IDP) characterized by the particularly low resonance dispersion, one has to rely on additional experiments at the expense of sensitivity loss, a significant increase of measurement time, and more complicated and tedious analyses.

Slow transverse relaxation of the $^{13}C^{\alpha}$ spins, which can be further decreased by deuteration, corresponds to the natural linewidth of 5–8 Hz even for relatively large protein systems. Unfortunately, the practical resolution in the HNCA spectra is usually almost ten times worse. Two main factors limit the resolution in the HNCA: (i) a large number of time-increments in the $^{13}C^{\alpha}$ dimension needed in the 3D experiment to achieve the high resolution. This leads to too long measurement time that can be unaffordable because of short sample stability and/or limitation on the measurement time at an NMR instrument; (ii) homonuclear one-bond coupling between $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ spins that produces a doublet with separation of approximately 35 Hz for every $^{13}C^{\alpha}$ signal and thus effectively broadens the spectral line. The former issue is well addressed by using fast pulsing[5,6] and non-uniform sampling (NUS) techniques.[7–12] A large number of methods for handling of the $^{1}J(C^{\alpha}–C^{\beta})$ coupling had been introduced over the last decades, including biochemical unlabelling of the β carbon atom to ^{12}C,[13–18] constant time evolution,[19] band-selective homonuclear decoupling,[20,21] and IPAP decoupling.[22,23] However, broad practical use of these techniques is hindered due to the inherent compromises in sensitivity, extra demands on sample isotope labelling, inability to deal with serine and threonine residues, and/or significant spectral distortions and artefacts.

A viable alternative to these experimental approaches is the virtual decoupling that is the post-acquisition deconvolution of

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

† Electronic supplementary information (ESI) available: Extended mathematical description of D-IRLS, sample preparation and experimental details, and results of simulations with injected peaks. See DOI: 10.1039/d0cc06188c

Open Access Article. Published on 28 October 2020. Downloaded on 6/4/2026 12:03:11 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Open Access Article. Published on 28 October 2020. Downloaded on 6/4/2026 12:03:11 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

the *J*-coupling, *i.e.* in- and anti-phase peak multiplets, at the signal processing stage.[24–29] The aim of this communication is to investigate the possibility of effective deconvolution in compressed sensing (CS) algorithms that are among the most powerful for the NUS spectra – to propose a method of selective deconvolution of individual spectral regions; and to demonstrate the relation of the deconvolution to the cornerstone CS concept of sparseness with the resulting benefits for the effectiveness of CS.

The fundamental relation between the NMR signal $f(t)$ detected in the time domain and the spectrum s is

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

where F is the measurement matrix composed of rows from the inverse Fourier transform matrix for every point in f. Thus, reconstruction of a spectrum from f reduces to solving the inverse linear system in eqn (1). For an undersampled (*i.e.* NUS) signal, the solution is not unique and additional constraints on the spectrum s are usually imposed. For example, using generalized Tikhonov regularization, the spectrum can be obtained as:

$$s = \arg\min_{x\in\mathbb{C}^n} \|Fx - f\|_Q{}^2 + \|x\|_D{}^2 \tag{2}$$

where, for a vector y and matrix G, $\|y\|_G{}^2$ denotes the weighted norm square $y^\dagger Gy$ with $y^\dagger$ denoting the conjugate transpose of y; $Q = \sigma^{-2}1$ is the inverse covariance matrix of noise in f, which is multiple of the identity matrix 1 and σ is the standard deviation of the noise; D is a diagonal matrix including weightings of the spectrum points and the Tikhonov regularization term. As will be shown below, Q is useful when dealing with the $^1J(C^\alpha–C^\beta)$ coupling, while matrix is the essential element of the Iterative Reweighted Least Squares (IRLS), one of the most popular algorithms for compressed sensing reconstruction of the NUS spectra (see ESI†).[30,31]

Assuming the same value of active $^1J(C^\alpha–C^\beta)$ coupling for all signals, the measured in experiment $^{13}C^\alpha$ signal f and the signal without the *J*-coupling $\tilde{f}$ are related as:

$$f = C\tilde{f} \text{ and } \tilde{f} = C^{-1}f \tag{3}$$

where C is a diagonal matrix with elements $\cos(\pi Jt)$ for every time point t in f.[32] If points in the measured signal f are corrupted by noise with inverse covariance matrix $Q = \sigma^{-2}1$, the noise in $\tilde{f}$ has the inverse covariance matrix $\tilde{Q} = \sigma^{-2}(CC)$. Then, the decoupled spectrum is

$$\tilde{s} = \arg\min_{x\in\mathbb{C}^n} \left\|Fx - \tilde{f}\right\|_{\tilde{Q}}{}^2 + \|x\|_D{}^2 \tag{4}$$

or equivalently (see ESI†),

$$\tilde{s} = \arg\min_{x\in\mathbb{C}^n} \|CFx - f\|_Q{}^2 + \|x\|_D{}^2. \tag{5}$$

The last equation shows that the post-acquisition deconvolution can be achieved in IRLS and any other algorithm based on equation akin to eqn (2), *e.g.* maximum entropy[24–27,29] and Multi-Dimensional Decomposition (MDD),[9] by using measurement matrix CF instead of F. Finally, we note that the deconvoluted spectrum contains half of the peaks relative to the undecoupled spectrum. Thus, it is sparser, and in accordance with the theory of compressed sensing,[33,34] it requires nearly half of the measured data points for successful reconstruction. This means that the virtual decoupling not only enhances spectral resolution but also provides conditions for higher quality CS reconstruction (see Theory in ESI†).

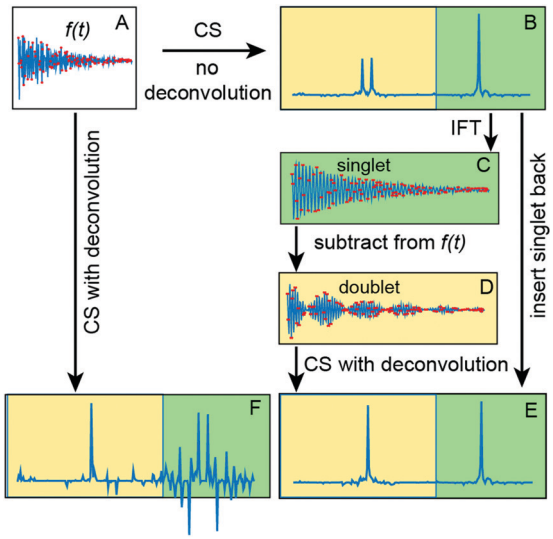


Fig. 1 Processing of a spectrum with region-selective deconvolution. (A) Measured time-domain signal that contains both a singlet and a doublet. IRLS reconstructions of (A) with and without deconvolution produce spectra (F and B), respectively. (C) The singlet (green) part of the spectrum (B) is converted back to the time-domain using inverse Fourier transform (IFT). (D) The original signal (A) after subtraction of (D). (E) The result of the region-selective deconvolution, *i.e.* combination of IRLS processing of (D) (yellow) and the green part of (B). Quality of both singlet and decoupled doublet signals in (E) is better than in (F).

Use of the deconvolution for the HNCA experiment is based on the assumption that $^1J(C^\alpha–C^\beta)$ coupling constants are nearly the same for all residues in the protein. The variation of the coupling values ±2.5 Hz[35] is lower than the line width determined by the transverse relaxation of $^{13}C^\alpha$ spins and, thus, does not pose a problem for the reconstruction (see ESI†). However, signals (singlets) from Gly residues that do not have C^β atoms, have no sparse representation in the columns of measurement matrix CF. Fig. 1 illustrates that this not only corrupts the Gly peaks in the deconvoluted spectrum but also affects other signals and reduces the overall quality of the reconstruction. To tackle this, we suggest a procedure of deconvolution-IRLS (D-IRLS) with the Gly-region selection as outlined in Fig. 1 (more details are found in the ESI†). We start with reconstructing the full undecoupled spectrum using matrix F. Because the $^{13}C^\alpha$ atoms usually have distinctly different chemical shifts with values lower than 45 ppm, we can subtract the well-reproduced signals in the Gly region from the original time-domain signal f, which is then used to reconstruct the spectrum with all signals except for Gly using eqn (5) with the measurement matrix CF. Finally, signals of Gly and other residues are combined into the full decoupled spectrum in the frequency domain.

Selection of the NUS acquisition schedule has a profound effect on the reconstruction quality. As f is multiplied by C^{-1} in

 This journal is © The Royal Society of Chemistry 2020

Open Access Article. Published on 28 October 2020. Downloaded on 6/4/2026 12:03:11 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

eqn (3), the noise is amplified the most for the points in $\tilde{f}$ at times, where $\cos(\pi Jt)$ function has small values (*i.e.* near $t = k/(2J)$, $k = 1, 3, 5$). In the weighted least squares method used to derive eqn (4), these points are used with low weights and thus contain relatively low information value. In the NUS schedule, it is logical to avoid these points and instead invest spectrometer time into more informative measurements. We used the signal amplitude matched NUS schedule with the sampling density corresponding to $|\cos(\pi Jt)|$ and rejecting points with probability less than 0.2[32,36] (Fig. S1, ESI†). Additionally, the schedule was in all cases relaxation-matched.

We demonstrate the new D-IRLS procedure using examples of two representative systems: intrinsically disordered human 441-residue Tau protein (the longest hTau40 isoform)[37] and the monomeric variant of the 509-residue globular photosensory module PAS-GAF-PHY of *Deinococcus radiodurans* phytochrome (*Dr*BphP$_{PSM}$).[38] For each protein, Fig. 2 shows the traditional low-resolution 3D HNCA spectrum superimposed with the resolution-enhanced spectrum obtained using D-IRLS with Gly-region selection. For *Dr*BphP$_{PSM}$ the two experiments were reconstructed using nearly the same number of NUS points corresponding to the same measurement time; for Tau, the low resolution experiment was around two times shorter. In the shown examples, the dramatically improved resolution of the D-IRLS spectrum allows us to observe sequential connectivities that are ambiguous in the traditional spectrum.

Fig. 2 and Fig. S2, S3 (ESI†) demonstrate that, in addition to the enhanced resolution, the D-IRLS spectra show higher or similar sensitivity in comparison to both the traditional low resolution and non-deconvoluted spectra. The peak connecting A87 and A88 in *Dr*BphP$_{PSM}$ spectra (Fig. 2B) provides a specific example of this. It is clearly seen in the 1D cross-sections in the D-IRLS spectrum. In the traditional experiment, the weak peak is completely masked by the slope of a stronger peak. In the non-deconvoluted spectrum, only one of the doublet components is present, which gives a completely wrong idea of the peak position.

Fig. S2 (ESI†) shows the $^{13}C/^{15}N$ projections from the spectra of both studied proteins, which confirms the superior quality of the spectra reconstructed with Gly-region selective D-IRLS. While the improved resolution in the spectra is anticipated from the deconvolution, the remarkable sensitivity of the D-IRLS spectrum can be explained by the increased sparsity favourable for the NUS reconstruction.

In order to extensively test the proposed D-IRLS method, we conducted simulations using synthetic peaks added to the 3D HNCA signal of Tau. Adding the simulated components with known positions and intensities to the time domain signal makes it possible to define the precision of the corresponding peak parameters derived from the reconstructed spectrum.[39] A detailed description of the simulations can be found in the ESI.† The results shown in Fig. S3 (ESI†) confirm that the peak intensities and positions are much more accurate when the D-IRLS deconvolution is augmented with the Gly-region selective procedure (Fig. 1E).

Notably, at low sampling levels (250 and 400 NUS points) the number of detected peaks from the lowest intensity fraction of

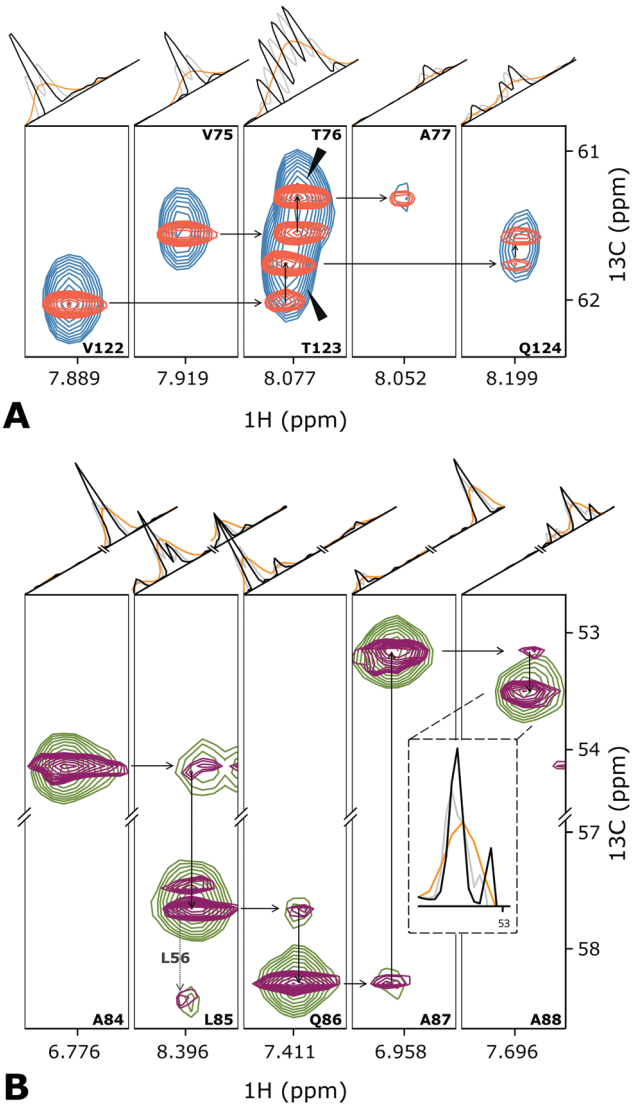


Fig. 2 Several planes from the 3D HNCA spectra of (A) Tau and (B) *Dr*BphP$_{PSM}$ showing the assignment walk for selected residues. Overlaid blue (green) and red (purple) contour levels depict traditional low-resolution and high-resolution spectra of Tau (*Dr*BphP$_{PSM}$) protein. For the peak annotations, we use the previously published assignment.[37,38] The one-dimensional cross-sections above the spectra planes are taken (orange) the low-resolution, (black) high-resolution deconvoluted, and (grey) high-resolution non-deconvoluted spectra.

the injected peaks is significantly larger in the deconvoluted spectrum in comparison to the IRLS without the deconvolution. This is fully in line with the theoretical consideration that the deconvoluted spectrum is much more sparse and thus can be successfully reconstructed with fewer measured points. In the wide range of NUS levels, cosine *J*-modulated sampling scheme provides spectra with comparable or somewhat better accuracy of the peak positions and intensities than the schedules matched to the exponential relaxation decay only. However, the main practical problem with the latter scheme is the necessity to increase and carefully adjust the Tikhonov regularization parameter λ in the IRLS algorithm, whereas the cosine-modulated sampling is much less demanding in this respect

Open Access Article. Published on 28 October 2020. Downloaded on 6/4/2026 12:03:11 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

and thus more robust. It is also worth noting that precision of the peak positions derived from the non-deconvoluted spectrum is somewhat better than in its decoupled counterpart provided that both components of the doublet are detected and resolved from other peaks.

In conclusion, we proposed an efficient CS method to improve the resolution, sensitivity and quality of NUS reconstruction using virtual decoupling at the processing stage. We presented a complete mathematical description of the spectrum deconvolution in terms of the generalized Tikhonov regularization formalism. We also showed that removing singlets from the spectrum before decoupling significantly improves results. The method was demonstrated on the 3D HNCA spectra of two large systems prototypical for the intrinsically disordered and globular proteins. The new CS virtual decoupling technique will enable the sequential signal assignment for many challenging proteins and will be useful for other types of NMR spectra in a variety of applications.

KK and PK thank the Foundation for Polish Science for support *via* the FIRST TEAM program co-financed by the European Union under the European Regional Development Fund no. (POIR.04.04.00-00-4343/17-00). BMB and SW acknowledge funding from the Knut och Alice Wallenberg Foundation. VO thanks for the support by the Swedish Research Council (Research Grant 2019-3661).

Conflicts of interest

There are no conflicts to declare.

References

1 M. Ikura, L. E. Kay and A. Bax, *Biochemistry*, 1990, **29**, 4659–4667.
2 L. E. Kay, M. Ikura, R. Tschudin and A. Bax, *J. Magn. Reson.*, 1990, **89**, 496–514.
3 M. Salzmann, K. Pervushin, G. Wider, H. Senn and K. Wüthrich, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 13585–13590.
4 Z. Solyom, M. Schwarten, L. Geist, R. Konrat, D. Willbold and B. Brutscher, *J. Biomol. NMR*, 2013, **55**, 311–321.
5 K. Pervushin, B. Vögeli and A. Eletsky, *J. Am. Chem. Soc.*, 2002, **124**, 12898–12902.
6 E. Lescop, P. Schanda and B. Brutscher, *J. Magn. Reson.*, 2007, **187**, 163–169.
7 J. C. J. Barna, E. D. Laue, M. R. Mayger, J. Skilling and S. J. P. Worrall, *J. Magn. Reson.*, 1987, **73**, 69–77.
8 J. C. Hoch, A. S. Stern and M. Mobli, *Maximum Entropy Reconstruction*, Update based on original article by Jeffrey C. Hoch, Encyclopedia of Magnetic Resonance, 1996, John Wiley & Sons Ltd., 2007, DOI: 10.1002/9780470034590.emrstm0299.pub2.
9 V. Orekhov and V. A. Jaravine, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2011, **59**, 271–292.
10 K. Kazimierczuk and V. Y. Orekhov, *Angew. Chem., Int. Ed.*, 2011, **50**, 5556–5559.
11 M. Bostock and D. Nietlispach, *Concepts Magn. Reson., Part A*, 2017, **46**, e21438.
12 S. Robson, H. Arthanari, S. G. Hyberts and G. Wagner, *Methods in Enzymology*, Elsevier, 2019, vol. 614, pp. 263–291.
13 D. M. LeMaster and D. M. Kushlan, *J. Am. Chem. Soc.*, 1996, **118**, 9255–9264.
14 P. E. Coughlin, F. E. Anderson, E. J. Oliver, J. M. Brown, S. W. Homans, S. Pollak and J. W. Lustbader, *J. Am. Chem. Soc.*, 1999, **121**, 11871–11874.
15 M. Kainosho, T. Torizawa, Y. Iwashita, T. Terauchi, A. M. Ono and P. Güntert, *Nature*, 2006, **440**, 52–57.
16 P. Lundström, K. Teilum, T. Carstensen, I. Bezsonova, S. Wiesner, D. F. Hansen, T. L. Religa, M. Akke and L. E. Kay, *J. Biomol. NMR*, 2007, **38**, 199–212.
17 K. Takeuchi, Z.-Y. J. Sun and G. Wagner, *J. Am. Chem. Soc.*, 2008, **130**, 17210–17211.
18 S. A. Robson, K. Takeuchi, A. Boeszoermenyi, P. W. Coote, A. Dubey, S. Hyberts, G. Wagner and H. Arthanari, *Nat. Commun.*, 2018, **9**, 356.
19 R. Powers, A. M. Gronenborn, G. M. Clore and A. Bax, *J. Magn. Reson.*, 1991, **94**, 209–213.
20 H. Matsuo, E. Kupče, H. Li and G. Wagner, *J. Magnet. Reson. Ser. B*, 1996, **113**, 91–96.
21 P. W. Coote, S. A. Robson, A. Dubey, A. Boeszoermenyi, M. Zhao, G. Wagner and H. Arthanari, *Nat. Commun.*, 2018, **9**, 3014.
22 P. Andersson, J. Weigelt and G. Otting, *J. Biomol. NMR*, 1998, **12**, 435–441.
23 M. Ottiger, F. Delaglio and A. Bax, *J. Magn. Reson.*, 1998, **131**, 373–378.
24 A. A. Bothner-By and J. Dadok, *J. Magn. Reson.*, 1987, **72**, 540–543.
25 M. A. Delsuc and G. C. Levy, *J. Magn. Reson.*, 1988, **76**, 306–315.
26 Z. Serber, C. Richter, D. Moskau, J.-M. Böhlen, T. Gerfin, D. Marek, M. Häberli, L. Baselgia, F. Laukien, A. S. Stern, J. C. Hoch and V. Dötsch, *J. Am. Chem. Soc.*, 2000, **122**, 3554–3555.
27 N. Shimba, A. S. Stern, C. S. Craik, J. C. Hoch and V. Dötsch, *J. Am. Chem. Soc.*, 2003, **125**, 2382–2383.
28 N. Shimba, H. Kovacs, A. S. Stern, A. M. Nomura, I. Shimada, J. C. Hoch, C. S. Craik and V. Dötsch, *J. Biomol. NMR*, 2004, **30**, 175–179.
29 R. Kerfah, O. Hamelin, J. Boisbouvier and D. Marion, *J. Biomol. NMR*, 2015, **63**, 389–402.
30 E. J. Candes, M. B. Wakin and S. P. Boyd, *J. Fourier Anal. Appl.*, 2008, **14**, 877–905.
31 K. Kazimierczuk and V. Y. Orekhov, *J. Magn. Reson.*, 2012, **223**, 1–10.
32 Other functional forms are possible for different multiplet types, *e.g.* sine-modulation for an anti-phase doublet.
33 E. Candes, J. Romberg and T. Tao, *IEEE Trans. Inf. Theory*, 2004, **52**, 489–509.
34 S. Foucart and H. Rauhut, *Bull. Am. Math.*, 2017, **54**, 151–165.
35 J. M. Schmidt, M. J. Howard, M. Maestr-Martínez, C. S. Pérez and F. Löhr, *Magn. Reson. Chem.*, 2009, **47**, 16–30.
36 V. Jaravine, I. Ibraghimov and V. Y. Orekhov, *Nat. Methods*, 2006, **3**, 605–607.
37 R. L. Narayanan, U. H. N. D urr, S. Bibow, J. Biernat, E. Mandelkow and M. Zweckstetter, *J. Am. Chem. Soc.*, 2010, **132**, 11906–11907.
38 E. Gustavsson, L. Isaksson, C. Persson, M. Mayzel, U. Brath, L. Vrhovac, J. A. Ihalainen, B. G. Karlsson, V. Orekhov and S. Westenhoff, *Biophys. J.*, 2020, **118**, 415–421.
39 M. A. Zambrello, A. D. Schuyler, M. W. Maciejewski, F. Delaglio, I. Bezsonova and J. C. Hoch, *Methods*, 2018, **138–139**, 62–68.

 This journal is © The Royal Society of Chemistry 2020