A partial sum of singular-value-based reconstruction method for non-uniformly sampled NMR spectroscopy

Zhangren Tu\textsuperscript{1,2}\dagger | Zi Wang\textsuperscript{3} | Jiaying Zhan\textsuperscript{1} | Yihui Huang\textsuperscript{3} | Xiaofeng Du\textsuperscript{1} | Min Xiao\textsuperscript{2} | Xiaobo Qu\textsuperscript{3} | Di Guo\textsuperscript{1}

\textsuperscript{1}School of Computer and Information Engineering, Xiamen University of Technology, Xiamen, People’s Republic of China
\textsuperscript{2}School of Opto-Electronic and Communication Engineering, Xiamen University of Technology, Xiamen, People’s Republic of China
\textsuperscript{3}Department of Electronic Science, Xiamen University, Xiamen, People’s Republic of China

Abstract
The nuclear magnetic resonance (NMR) spectroscopy has fruitful applications in chemistry, biology and life sciences, but suffers from long acquisition time. Non-uniform sampling is a typical fast NMR method by undersampling the time-domain data of the spectrum but need to restore the fully sampled data with proper constraints. The state-of-the-art method is to model the time-domain data as the sum of exponential functions and reconstruct these data by enforcing the low rankness of Hankel matrix. However, this method is solved by minimizing the sum of singular values of the Hankel matrix, which leads to the distortion of low-intensity spectral peaks. Here, a low rank Hankel matrix reconstruction approach with a partial sum of singular values is proposed to protect small singular values, which can faithfully reconstruct all peaks. Results on both synthetic and realistic NMR spectroscopy show that the proposed method can reconstruct a more consistent spectrum to the fully sampled one than other state-of-the-art methods and have particular advantages on preserving low-intensity peaks.

1 | INTRODUCTION

Nuclear magnetic resonance (NMR) spectroscopy provides fruitful atomic-level information on molecular structures. It is an indispensable tool for the analysis in physics, chemistry [1], structural biology [2,3], medical science and other fields [4]. The data acquisition time of NMR spectra increases rapidly with dimensions and spectral resolutions [5]. Non-uniform sampling (NUS) [6–10] is commonly used to accelerate the data acquisition by undersampling the time-domain data of the spectrum. More recently, the deep learning approach has been proposed to accelerate NMR by using solely synthetic NMR signals, which lifts the prohibiting demand for a large volume of realistic training data usually required for a deep learning approach [11]. The deep learning NMR, however, requires the huge database and long time to train the neural network, although the spectra reconstruction from NUS is ultrafast.

Modern methods reconstruct high-quality NMR spectra from NUS data by imposing appropriate constraints on it. Typical constraints include sparsity of the spectrum [12–16] and low rankness of time-domain data, which is called the free induction decay (FID) signal, of the spectrum [17–22]. The former assumes that non-zero values are rare, for example, sparsity [12–16], in the spectrum and explores the $\ell_1$ norm [23–25] to constrain spectral sparsity. Although narrow peaks can be reconstructed very well, reconstruction of broad peaks may be compromised since they violate the assumption of sparsity. The latter [17] explores the low-rank Hankel matrix (LRHM) constructed from the FID assuming that the number of peaks is much fewer than the length of the spectrum. Due to the independence of the rank on the peak width, the LRHM method can simultaneously reconstruct both narrow and broad peaks, and has a significant advantage on preserving low-intensity broad peaks [17]. More advances of LRHM can be found in a recent review [26].

\dagger Grants/Award Number: 61871341, 61672335, Fujian Province of China, Grant/Award Number: 2019J01855

Correspondence
Di Guo, School of Computer and Information Engineering, Xiamen University of Technology, Xiamen 361024, China.
Email: guodi@xmut.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 61871341, 61672335, 61971361, U1805264; Natural Science Foundation of Fujian Province of China, Grant/Award Number: 2019J01855
In the LRHM method [17], the FID signal is transformed into a Hankel matrix, whose rank is equal to the number of spectral peaks and supposed to be small. By introducing a nuclear norm, defined as the sum of singular values, as an approximation of rank, the FID is reconstructed by finding the low-rank solution subject to the acquired data. To achieve a better performance on recovering the low-intensity peaks, the weighted sum of singular values [19] is introduced as a better approximation of rank to LRHM, but the improvement is limited and not robust to sampling trials (See Section 4.1).

In this treatise, a new LRHM approach is proposed based on the partial sum of singular values (PSSV). This method tries to protect small singular values and introduces a better strategy to approximate the rank of Hankel matrix. Results on the synthetic and realistic NMR data show that the proposed method outperforms the state-of-the-art low-rank methods on preserving low-intensity peaks and provides more robust reconstructions.

**TABLE 1** The reconstruction algorithm of the proposed method

| Initialization: Input y, U, R, set maximal β = 10^3, convergence condition η_{hi} = 10^{-5}, and maximal number of iteration K = 10^3. Initialize the solution x^0 = U^T y and β = 1, the number of iterations k = 0 and η_k = 1. |
|-----------------------------------------------|
| **Main:** |
| While β_j < β_{min} and η_{j+1} < η_{hi}, do |
| | 1) Solve the minimizer of Equation (8) as follows: |
| | While (η_k ≥ η_{hi}) or (k < K), do |
| | a) Update x^{k+1} according to Equation (11); |
| | b) Update Z^{k+1} according to Equation (13); |
| | c) Compute η_{k+1} = ∥x^{k+1} − x^{k}∥ / ∥x^{k}∥ and k ← k + 1; |
| | End while |
| | 2) Update x ← x^{k+1} and x ← x; |
| | 3) Set k = 0, η_k = 1 and β_j+1 ← 2β_j; |
| | End while |

| Output: The reconstructed FID signal ̂x. |
2 | RELATED WORK

In NMR spectroscopy, the time-domain FID signal is commonly modelled as the sum of exponential functions according to [10,11,17–22,26–28]

\[ x_n = \sum_{j=1}^{J} (a_j e^{i \phi_j}) e^{-\frac{\sigma_j}{2} e^{i n \tau_j}}, \quad n \in \{1, 2, \ldots, N\}, \]  

where \( \Delta t \) is the time interval; \( J \) is the number of spectral peaks; \( a_j, \phi_j, \tau_j \) and \( \sigma_j \) are the amplitude, phase, damping factor and the frequency of the \( j \)th spectral peak, respectively.

In NUS, an unduplicated \( n \) was chosen from partial entries in the index \( \{1, 2, \ldots, N\} \) where \( N \) is the total number of fully sampled data points. Assuming the total number of \( n \) is \( M \), then the sampling rate \( g \) is defined as \( g = M/N \). The sampled data \( y \in \mathbb{C}^M \) in NUS is defined as

\[ y = Ux + e, \]

where \( U \in \mathbb{C}^{M \times N} \) is an undersampling operator that defines how to acquire the FID with NUS and \( e \in \mathbb{C}^M \) is the additive Gaussian noise, which is the most common noise in NMR spectroscopy [27] and has been widely applied in NMR spectra reconstruction [10,11,14,17,20,22].

In the LRHM method, the reconstruction of FID signal \( x \) is formulated as a nuclear norm minimization problem of the Hankel matrix according to [17]

\[ \min_{x} \|Rx\|_\omega + \lambda \|y - Ux\|_2^2, \]  

where \( R \) denotes a Hankel operator converting the \( x \) into a Hankel matrix \( Rx \), \( \|\cdot\|_\omega \) represents the nuclear norm (sum of singular values) [29,30], \( \|\cdot\|_2 \) represents the \( L_2 \) norm, and \( \lambda \) is a regularization parameter that balances the low rankness and the data consistency. To distinguish this basic model in Equation (3) from other modifications discussed later, we call this method as nuclear norm-based LRHM (NN-LRHM).

However, the NN-LRHM provides suboptimal reconstruction when NUS data are few in highly accelerated NMR. Since large singular values usually submerge the small ones in NUS and the amplitude of singular value reflects the intensity of peaks, the reconstruction of low-intensity peaks becomes hard and one may observe the distortion or even disappearance of these weak peaks [19].

To achieve better performance, the low-rank constraint is enforced by minimizing the weighted nuclear norm of the LHRM (WNN-LRHM) [19]. Its core idea is to better approximate the low rankness by assigning a big (or small) weight on the small (or big) singular value [19]. The model is as follows:

\[ \min_{x} \|Rx\|_\sigma + \lambda \|y - Ux\|_2^2, \]

where

\[ \|Rx\|_\omega = \sum_{i=1}^{S} w_i \sigma_i, \]

where \( w = [w_1, \ldots, w_i, \ldots, w_S]^T \) includes a weight \( w_i \) for the \( i \)th singular value \( \sigma_i \).

Although the NMR reconstruction performance of WNN-LRHM is greatly improved, particularly on low-intensity peaks, this improvement is not stable due to the incorrect learning of weights from typical nuclear norm reconstruction (see Section 4.1).

3 | PROPOSED METHOD

To better recover a general low-rank matrix, a new metric called PSSV was proposed to outperform the nuclear norm in computer vision [31]. One main advantage of PSSV is that it avoids simultaneously minimizing all singular values and provides a
better approximation to matrix rank. The PSSV has never been extended into the LRHM reconstruction in NMR spectroscopy.

In this work, we propose an LRHM reconstruction model by minimizing the PSSV, which is called PSSV-LRHM, as follows:

$$\min \|Rx\|_* - \|P_r(Rx)\|_* + \frac{\lambda}{2}||y - Ux||_2^2,$$  

(6)

where $P_r(\cdot)$ is the matrix projection operator to the matrix of rank $r$, and it is defined as

$$P_r(X) = H_{1,r}^*XV_{1,r},$$  

(7)

where $H_{1,r}$ and $V_{1,r}$ are the matrices consisting of the singular vectors corresponding to the $r$ largest singular values of $X$. The PSSV minimizes the sum of small singular values by subtracts the sum of large singular values $\|P_r(Rx)\|_*$ from all singular values $\|Rx\|_*$, leading to protection of small singular values in the reconstruction. Since the intensity of spectral peaks is basically reflected by the singular values, preserving the low-intensity peaks is expected in PSSV-LRHM.

To efficiently solve the proposed PSSV-LRHM model, the alternating direction minimization with continuation algorithm [32] is modified to fit the Hankel matrix reconstruction problem. The main steps are derived as follows.

The augmented Lagrange of Equation (6) is

$$G(x, Z, \beta) = \|Z\|_r + \frac{\beta}{2}\|Rx - Z\|_F^2 + \frac{\lambda}{2}\|y - Ux\|_2^2,$$  

(8)

where $\beta$ is a penalty variable; a larger $\beta$ leads to better approximation to the rank. $\|\cdot\|_r$ represents Frobenius norm and $\|\cdot\|_*$ represents the PSSV with targeted rank $r$ [31].

The algorithm consists of two loops of iterations. In the outer loop, we progressively increase $\beta$. In the inner loop, we alternatively solve Equation (8) by two sub-problems:

$$\left\{\begin{array}{l}
x^{k+1} = \arg\min_G(x, Z^k, \beta) \\
Z^{k+1} = \arg\min_G(x^{k+1}, Z, \beta)
\end{array}\right..$$  

(9)

(1) Fixing $Z^k$, $x^{k+1}$ is obtained by solving

$$\min \frac{\beta}{2}\|Rx - Z^k\|_F^2 + \frac{\lambda}{2}\|y - Ux\|_2^2,$$  

(10)

whose solution is

$$x^{k+1} = \left(\lambda U^TU + \beta R^TR\right)^{-1}\left(\lambda U^Ty + \beta R^TZ^k\right).$$  

(11)

(2) Fixing $x^{k+1}$, $Z^{k+1}$ is obtained by solving

$$\min \frac{\beta}{2}\|Z\|_r + \frac{\lambda}{2}\|Rx^{k+1} - Z\|_F^2.$$

(12)

Here, a partial singular value thresholding operator [31] is used to solve Equation (12), whose solution is
The algorithm converges after \( k = \min(k_{\text{max}}, K) \) pseudocodes.

The pseudo-code of this reconstruction algorithm is shown in Table 1.

4 | EXPERIMENTAL RESULTS

The proposed method is compared with two state-of-the-art NMR reconstruction methods, including NN-LRHM and WNN-LRHM. Their performances are tested on both synthetic and realistic NMR data. All the 1D and 2D spectra reconstruction were performed using MATLAB 2017b (Mathworks Inc., Natick, MA) on a personal computer with 1.80 GHz 8 core CPU and 8 GB RAM.
Figure 6: The zoom-in region for the reconstruction of the 2D HSQC spectrum. (a)–(d) are zoom-in regions of (a)–(d) in Figure 5, respectively. The location of this region is marked with a dash line in Figure 5a. Heteronuclear single-quantum correlation spectroscopy (HSQC) plays a crucial role in identifying the amino acid sequences and their secondary structures.

Figure 7: Peak intensity correlation between fully sampled spectrum and reconstructed spectrum on the 2D HSQC of cytosolic CD79b protein. (a)–(c) are correlation plots of all peaks using nuclear norm-based low-rank Hankel matrix (LRHM), weighted NN-based LRHM and partial sum of singular values LRHM, respectively. The $R^2$ denotes the square of Pearson correlation coefficient; the closer the value is to 1, the closer the reconstructed spectrum is to the reference. HSQC, Heteronuclear single-quantum correlation spectroscopy.

Figure 8: Peak low-intensity correlation between fully sampled spectrum and reconstructed spectrum on the 2D HSQC of cytosolic CD79b protein. (a)–(c) are correlation plots of low-intensity peaks using nuclear norm-based low-rank Hankel matrix (LRHM), weighted nuclear norm-based LRHM and partial sum of singular values LRHM, respectively. The $R^2$ denotes the square of Pearson correlation coefficient; the closer the value is to 1, the closer the reconstructed spectrum is to the reference. HSQC, Heteronuclear single-quantum correlation spectroscopy.

4.1 Synthetic data

Two synthetic spectra are generated according to the exponential function in Equation (1) and their parameters are listed in Tables A1 and A2, respectively. For the first spectrum (Figure 1a), the broadness of peak 1, 2, 3, 4 and 5 are decreased and their intensity is increased accordingly since a larger $\tau$ results in a slower signal attenuation. Therefore, the broadest peak, which is also the lowest one, is peak 1 as shown in Figure 1a. The second spectrum (Figure 4a) contains dense peaks with different intensities. The total number of data points for each synthetic data is 512. Then, Gaussian noise [10,11,14,17,20,22,27] with zero mean and standard deviation of 0.005 is added to the two spectra. Finally, FID are sampled according to Poisson-gap NUS [33].

In general, the larger number of peaks, the greater the number of samples should be. Our empirical observations indicate that a faithful reconstruction usually requires the number of samples $M$ being larger than five times of the number of peaks $J$, that is, $M > 5J$. 
Reconstructions on the first spectrum are shown in Figure 1. With very limited 51 sampled points under the sampling rate 10%, the height of moderate-intensity peak (peak 3 in Figure 1b) is incorrect and low-intensity peaks (peaks 1 and 2 in Figure 1b) are seriously distorted by NN-LRHM. By learning the weights on singular values from Figure 1b, moderate-intensity peaks (peak 3 in Figure 1c) are recovered much better by WNN-LRHM, but the peak shape or intensity of low-intensity peaks (peaks 1 and 2 in Figure 1c) is still incorrect. The proposed PSSV-LRHM remarkably improves the broadest and lowest intensity peak (peak 1 in Figure 1d) than other methods and obtains the most consistent spectrum to the ground truth (Figure 1a).

To test the robustness to randomness in NUS, we compare the quantitative analysis on the spectrum intensities correlation over 100 sampling trials. As shown in Figure 2, PSSV-LRHM achieves the highest correlations and the lowest standard deviations of correlations on low-intensity peaks (peaks 1 and 2) and comparable performance on other peaks. The scatter plot of correlations (Figure 3) for low-intensity peaks imply that, in most cases, PSSV-LRHM obtains higher correlations than other methods. Even when WNN-LRHM gets extremely low correlation that is close to 0.2, the new method obtains correlation that is close to 1.

For the reconstruction of the dense spectrum in Figure 4, the proposed approach obtains more accurate peak intensity than other methods and more consistent spectral shape to the ground truth spectrum (Figure 4a). Thus, the PSSV-LRHM achieves the best reconstruction among three methods.
4.2 | NMR data

The realistic 2D NMR data include a $^1$H–$^{15}$N Heteronuclear single-quantum correlation (HSQC) spectrum of cytosolic CD79b protein [17] and a $^1$H–$^{15}$N HSQC spectrum of ubiquitin [11]. For realistic 2D NMR, the undersampling is only performed on the indirect dimension, since the data acquisition in this dimension is time consuming due to the evolution of magnetization in NMR physics [13]. Another dimension called direct dimension is fully sampled since the data acquisition is fast.

As shown in Figure 5a, the 2D $^1$H–$^{15}$N HSQC spectrum of cytosolic CD79b protein is acquired for 300 μM $^{15}$N–$^{13}$C labelled sample of cytosolic CD79b in 20 mM sodium phosphate buffer, pH 6.7 at 25°C on 800 MHz Bruker AVANCE III HD spectrometer equipped with 3-mm TCI cryoprobe. The fully sampled spectrum consists of 1024 data points in the direct dimension ($^1$H) and 256 data points in the indirect dimension ($^{15}$N).

The reconstructed 2D NMR spectra are shown in Figure 5. Some missing peaks (Figures 5b,c and 6) are presented in the reconstruction obtained by NN-LRHM and WNN-LRHM, while these distortions are not observed in the PSSV-LRHM reconstruction. The quantitative analysis on the peak intensities correlation in Figures 7 and 8 indicate that the proposed method achieves the highest correlations among three methods.

As shown in Figure 9a, the 2D $^1$H–$^{15}$N HSQC spectrum of ubiquitin is acquired at 298.2 K temperature on an 800-MHz Bruker spectrometer. The fully sampled spectrum consists of 1024 data points in the direct dimension ($^1$H) and 98 data points in the indirect dimension ($^{15}$N). According to the Poisson-gap NUS, 25% of both data are non-uniformly sampled in the indirect dimension [33].

The reconstructed 2D NMR spectra are shown in Figure 9. Some pseudo peaks (Figure 9b,c) are presented in the reconstruction obtained by NN-LRHM and WNN-LRHM, while these distortions are not observed in the PSSV-LRHM reconstruction. The quantitative analysis on the peak intensities correlation in Figures 10 and 11 indicate that the proposed method achieves the highest correlations among three methods.

These observations imply that the PSSV-LRHM can reconstruct the 2D NMR spectra more faithfully, especially for low-intensity peaks.
5 | DISCUSSIONS

To verify the performances of the proposed method and other two methods, their computational time, the reconstruction ability of the minimum low-intensity peak and high dimensional spectra are compared here. Besides, the limitation of the proposed method is discussed.

5.1 | Computational time

The time taken to reconstruct 2D HSQC spectrum of cytosolic CD79b protein by NN-LRHM, WNN-LRHM and PSSV-LRHM are 35.1, 167.8 and 212.6 s, respectively. The reconstruction time of HSQC spectrum of ubiquitin by NN-LRHM, WNN-LRHM and PSSV-LRHM are 38.5, 169.7 and 174.9 s, respectively. The time cost of the proposed method is 4–6 times of the basic NN-LRHM and is 1–1.3 times of the state-of-the-art approach, WNN-LRHM. Thus, it is valuable to adopt the new method in NMR since much better performance is achieved with comparable time of the state-of-the-art method.

5.2 | Minimum low-intensity peak

In NMR, the low-intensity peak usually means the intensity of the lowest peak is much lower than that of the highest one. To verify the performance of minimum intensity in reconstruction, a high dynamic range spectrum shown in Figure 12a is generated, whose lowest intensity is $10^{-3}$ of the highest one. The synthetic spectra are generated according to exponential functions in Equation (1) and their parameters are listed in Table A3.

Results show that, when the sampling rate is 10%, the proposed method (Figure 13d) holds the advantage over NN-LRHM (Figure 13b) and WNN-LRHM (Figure 13c) on the reconstruction of the minimum low-intensity peak (marked with an arrow in Figure 12 and its zoom-in in Figure 13). The intensity of the lowest peak in NN-LRHM, however, is relatively lower than that of fully sampled spectra (Figure 13a).

By increasing the sampled rate to 25% (Figures 14 and 15), high-fidelity reconstruction of the minimum low-intensity peak (Figure 16d) can be achieved while compared methods still introduce strong artefacts (Figure 15b,c).
Thus, the proposed provide much better reconstruction of the minimum low-intensity peak whose intensity is $10^{-3}$ of the highest one.

### 5.3 High-dimension spectra

We add the reconstruction of a synthetic 3D spectrum. The spectrum is with a size $20 \times 20 \times 20$ are obtained by performing the Fourier transform on the 3D FID generated according to Equation (18) with parameters listed in Table A4.

The fully sampled spectrum is shown in Figure 16a. The lowest intensity peak, which is marked with an arrow, is reconstructed much better by the proposed method (Figure 16d) than NN-LRHM (Figure 16b) and WNN-LRHM (Figure 16c). The advantage of the PSSV is also confirmed by the lowest reconstruction error, relative $l_2$ norm error (RLNE), under 10 trials of NUS. Thus, the proposed method can be extended to high-dimensional spectrum reconstruction and still outperforms other methods on achieving lower reconstruction error and preserving low-intensity signals.

The current PSSV-LRHM as well as other LRHM methods work for complex spectra reconstruction but cannot be applied to realistic high-dimensional spectroscopy, such as 3D NMR [10, 11, 34], since they cannot handle the hypercomplex format in the indirect dimension [35]. We remark this as a valuable future work.

### 5.4 Limitation by noise

The limitation of the proposed method is the same as other methods. When the noise is relatively high, the lowest peaks may be distorted by all methods, although all other higher-intensity
peaks are reconstructed much better (Figure 17). In addition, the proposed PSSV obtains slightly higher peak correlation for all peaks than other methods (Figure 18).

6 | CONCLUSION

An LRHM method with a PSSV is proposed to reconstruct the undersampled data in accelerated NMR spectroscopy. Experimental results on both synthetic and realistic NMR data demonstrate that the new approach reconstructs spectral peaks more faithfully and robustly, particularly for low-intensity peaks. Possible future work could be extended into higher dimensional NMR, improving the reconstruction under high noise, and incorporating the protection of weak signals into exponential-based [17] and LRHM-based deep learning reconstruction [36], which may enable ultrafast and high-fidelity reconstruction of fast NMR [17,36,37].

ACKNOWLEDGMENTS

The work was supported by the National Natural Science Foundation of China (61871341, 61672335, 61971361, U1805264), and Natural Science Foundation of Fujian Province of China (2019J01855). The authors thank Vladislav Orekhov for sharing the NMR data used in paper [17], and thank reviewers and editors for valuable suggestions.

ORCID

Zhangren Tu https://orcid.org/0000-0002-0863-2454

REFERENCES

1. Zhdanko, A., Ströbele, M., Maier, M.E.: Coordination chemistry of gold catalysts in solution: a detailed NMR study. Chem. Eur. J. 18(46), 14732–14744 (2012)
2. Betz, M., Saxena, K., Schwalbe, H.: Biomolecular NMR: a chaperone to drug discovery. Curr. Opin. Chem. Biol. 10(3), 219–225 (2006)
3. Cavalli, A., et al.: Protein structure determination from NMR chemical shifts. Proc. Natl. Acad. Sci. U. S. A. 104(23), 9615–9620 (2007)
4. Tkáč, I., et al.: In vivo 1H NMR spectroscopy of the human brain at high magnetic fields: metabolite quantification at 4T Vs. 7T. Magn. Reson. Med. 62(4), 868–879 (2009)
5. Jaravel, V., Ibragimov, I., Orekhov, V.Y.: Removal of a time barrier for high-resolution multidimensional NMR spectroscopy. Nat. Methods. 3(8), 605–607 (2006)

6. Barna, J., et al.: Exponential sampling, an alternative method for sampling in two-dimensional NMR experiments. J. Magn. Reson. 73(1), 69–77 (1987)

7. Jiang, B., et al.: Gridling and fast fourier transformation on non-uniformly sampled multidimensional NMR data. J. Magn. Reson. 204(1), 165–168 (2010)

8. Lafon, O., et al.: Fast and high-resolution stereochemical analysis by nonuniform sampling and covariance processing of anisotropic natural abundance 2D 2H NMR datasets. Chem. Eur. J. 17(24), 6716–6724 (2011)

9. Tu, Z., et al.: A fast self-learning subspace reconstruction method for non-uniformly sampled nuclear magnetic resonance spectroscopy. Appl. Sci. 10(11), 3939 (2020)

10. Ying, J., et al.: Hankel matrix nuclear norm regularized tensor completion. IEEE Trans. Sig. Proc. 65(14), 3702–3717 (2017)

11. Qu, X., et al.: Accelerated nuclear magnetic resonance spectroscopy with low rank Hankel matrices and tensors. Magn. Reson. 209 (2), 352–358 (2011)

12. Qu, X., et al.: Fast sparse reconstruction algorithm for multidimensional NMR experiments. J. Magn. Reson. 73 (1), 69–77 (1987)

13. Qu, X., et al.: Reconstruction of self-sparseness reconstruction for N-dimensional exponential signals. IEEE Trans. Sig. Proc. 65(14), 3702–3717 (2017)

14. Guo, D., Lu, H., Qu, X.: A fast low rank Hankel matrix factorization algorithm and applications. IEEE Trans. Patt. Anal. 38(4), 744–758 (2015)

15. Jin, Z.-F., et al.: An alternating direction method with continuation for nonconvex low rank minimization. J. Sci. Comput. 66(2), 849–869 (2016)

16. Hyberts, S.G., Takeuchi, K., Wagner, G.: Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data. J. Am. Chem. Soc. 132(7), 2145–2147 (2010)

17. Qu, X., et al.: Solution structure of the second RRM domain of RBM5 and its unusual binding characters for different RNA targets. Biochemistry. 51(33), 6667–6678 (2012)

18. Bostock, M.J., Holland, D.J., Nietlispach, D.: Improving resolution in multidimensional NMR using random quadrature detection with compressed sensing reconstruction. J. Biomol. NMR. 68(2), 67–77 (2017)

19. Zhang, X., et al.: Image reconstruction with low-rankness and self-consistency of k-space data in parallel MRI. Med. Image Anal. 63, 101687 (2020)

20. How to cite this article: Tu Z, Wang Z, Zhan J, et al. A partial sum of singular-value-based reconstruction method for non-uniformly sampled NMR spectroscopy. IET Signal Process. 2021;15:14–27. https://doi.org/10.1049/si.2020.0456

APPENDIX

The ppm denotes parts per million by frequency, the unit of chemical shift. The chemical shift is defined as [20]:

$$\text{Chemical shift (ppm)} = \frac{f_{\text{sample}} - f_{\text{ref}}}{f_{\text{spec}}} \times 10^6$$

where $f_{\text{sample}}$ denotes resonance frequency of the sample, $f_{\text{ref}}$ the absolute resonance frequency of a standard compound measured in the same magnetic field and $f_{\text{spec}}$ the frequency of magnetic field strength of spectrometer.

The synthetic 3D signal is generated in the time domain as follows [10,27]:

$$y_{i_1,i_2,i_3} = \sum_{n=1}^{R} a_n \left(e^{-1/T_{1,n}+2\pi f_{1,n}}ight)^{i_1} \left(e^{-1/T_{2,n}+2\pi f_{2,n}}ight)^{i_2} \left(e^{-1/T_{3,n}+2\pi f_{3,n}}ight)^{i_3}$$

where $R$ is the number of exponentials (i.e. spectral peaks), the frequency $f_{n,n}$ is uniform randomly drawn from the interval $[0,1)$, $a_n$ is the amplitude, $T_{n,n}$ is the damping factor, and $i_n = 1, 2, \cdots, 20 (n = 1, 2, 3)$. The NUS is performed on the indirect plane that is constituted by the variables $i_1$ and $i_2$ [18,38].
**TABLE A1** Synthetic data for five peaks in Figure 1–3

| Peaks ID | Parameters |
|---------|------------|
|         | 1   | 2   | 3   | 4   | 5   |
| Amplitude | 0.3 | 0.4 | 0.5 | 1   | 1   |
| Damping factor | 0.01 | 0.02 | 0.03 | 0.04 | 0.08 |

Note: The broadest and the lowest peak is Peak 1.

**TABLE A2** Synthetic data for dense peaks in Figure 4

| Peaks ID | Parameters |
|---------|------------|
|         | Amplitude | Damping factor |
| 1       | 0.20      | 0.15          |
| 2       | 0.28      | 0.15          |
| 3       | 0.35      | 0.16          |
| 4       | 0.50      | 0.17          |
| 5       | 0.55      | 0.17          |
| 6       | 0.58      | 0.17          |
| 7       | 0.69      | 0.18          |
| 8       | 0.87      | 0.19          |
| 9       | 0.90      | 0.16          |
| 10      | 0.92      | 0.17          |
| 11      | 0.95      | 0.24          |
| 12      | 0.97      | 0.24          |
| 13      | 0.98      | 0.21          |
| 14      | 1.00      | 0.25          |

**TABLE A3** Parameters of high dynamic range spectrum in Figure 12

| Parameters | Peaks ID |
|------------|----------|
|            | 1 | 2 | 3 | 4 |
| Amplitude  | 0.001 0.01 0.1 1 |
| Damping factor | 0.05 0.05 0.05 0.05 |

**TABLE A4** Parameters of synthetic 3D spectrum in Figure 16

| Peaks ID | Parameters |
|----------|------------|
|          | Amplitude | Damping factor |
| 1        | 1         | 20            |
| 2        | 10        | 20            |
| 3        | 20        | 20            |
| 4        | 30        | 20            |
| 5        | 50        | 20            |
| 6        | 60        | 20            |
| 7        | 70        | 20            |
| Peaks ID | Parameters |
|----------|------------|
|          | Amplitude  | Damping factor |
| 8        | 80         | 20             |
| 9        | 90         | 20             |
| 10       | 100        | 20             |

**Table A4** (Continued)