

PCCP

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

COMMUNICATION

Non-uniform Sampling in EPR – optimizing data acquisition for HYSORE spectroscopy

Cite this: DOI: 10.1039/x0xx00000x

K. Kanth,^a Y. A. Tesiram,^a I. M. Brereton,^a M. Mobli*^a and J. R. Harmer*^aReceived 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

Non-uniform sampling combined with maximum entropy reconstruction is a powerful technique used in multi-dimensional NMR spectroscopy to reduce sample measurement time. We adapted this technique to the pulse EPR experiment Hyperfine Sublevel Correlation (HYSORE) and show that experimental times can be shortened by approximately an order of magnitude as compared to conventional linear sampling with negligible loss of information.

The experimental time required to record a multi-dimensional NMR or EPR spectrum is directly proportional to the number of points sampled in the indirect dimensions. In many two-dimensional (2D) pulse EPR experiments, such as HYSORE¹ (Fig. 1) or 2D DQ EPR,² data is collected with constant time spacing and point-by-point along both dimensions by recording the electron spin echo intensity at the end of the pulse sequence. Thus both dimensions of the experiment are indirect dimensions, i.e. an FID is not recorded along any dimension. This means that a significant amount of time is needed to carry out many 2D pulse EPR experiments, and 3D experiments are essentially unfeasible.

The same problem exists in multi-dimensional NMR experiments. For example 3D NMR experiments typically have two indirect dimensions and the measurement time with a linear sampling scheme limits the sensitivity and resolution of the data that can be collected in a reasonable time. To overcome this limitation the method of non-uniform sampling (NUS) combined with methods of spectrum reconstruction have been developed to a high level of sophistication.^{3,4,5,6}

So far, this approach has not been applied to multi-dimensional pulse EPR experiments. In this contribution, we present an analysis of two sets of NUS HYSORE data their spectrum reconstruction using a maximum entropy (MaxEnt)⁷ algorithm, and characterise the potential time savings and advantages of the method.

Measurement time in HYSORE studies is typically a limiting factor, particularly as the experiment needs to be repeated with different τ values to avoid suppressing peaks from the τ -dependent spectral blind spots, and is often carried out on a paramagnetic centre with a broad field-sweep EPR spectrum so that a set of HYSORE measurements are required at different field positions B_0 to determine the interactions, e.g. the relative orientation and principal values of the hyperfine and nuclear quadrupole couplings from magnetic nuclei coupled to the electron spin.

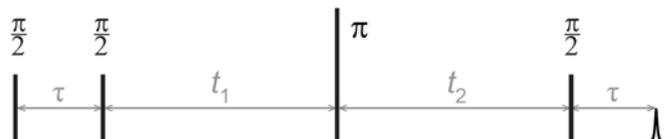


Fig. 1: Hyperfine sublevel correlation (HYSORE) pulse sequence, the time domain matrix is collected as $Y(t_1, t_2)$ for a fixed τ . For an $S = 1/2$ spin the experiment correlates nuclear frequencies in the two electron spin manifolds, yielding cross-peaks symmetric to the diagonal in the frequency domain spectrum. Unwanted blind spots occur at $\nu_{\text{blind}} = n/\tau$; $n = 0, 1, 2, \dots$

We will use two distinctly different sets of HYSORE data in our analysis, one recorded at X-band (9.7822 GHz) and reporting on the ¹H couplings to the paramagnetic [2Fe-2S]⁺ cluster in palustrisredoxin-B (PuxB) from *Rhodospseudomonas palustris* CGA009 PuR,⁸ and the second recorded at Q-band (35.309 GHz) and reporting on the three ¹⁴N couplings from the stable aminyl radical metal complex [Rh(trop2N)(bipy)]OTf⁻ in THF-acetone⁹. These data are shown in Fig. 2 and 4, respectively. The time-domain data (Fig. 2-A1 and 4-A1) were acquired with linear sampling and with a spectral bandwidth ($1/\Delta T$) sufficient to avoid peak aliasing. The corresponding absolute-value frequency spectra shown in Fig. 2-A2 and 4-A2 were computed from a 2D discrete Fourier transformation (DFT) of the time-domain data after a baseline correction to remove the unmodulated contributions, apodisation with a Gaussian window, and zero filling to 1024 points.

The signal-to-noise ratios (SNR) are $SNR \sim 70$ ($[2Fe-2S]^+$) and $SNR \sim 40$ ($[Rh(\text{trop}2N)(\text{bipy})]$).

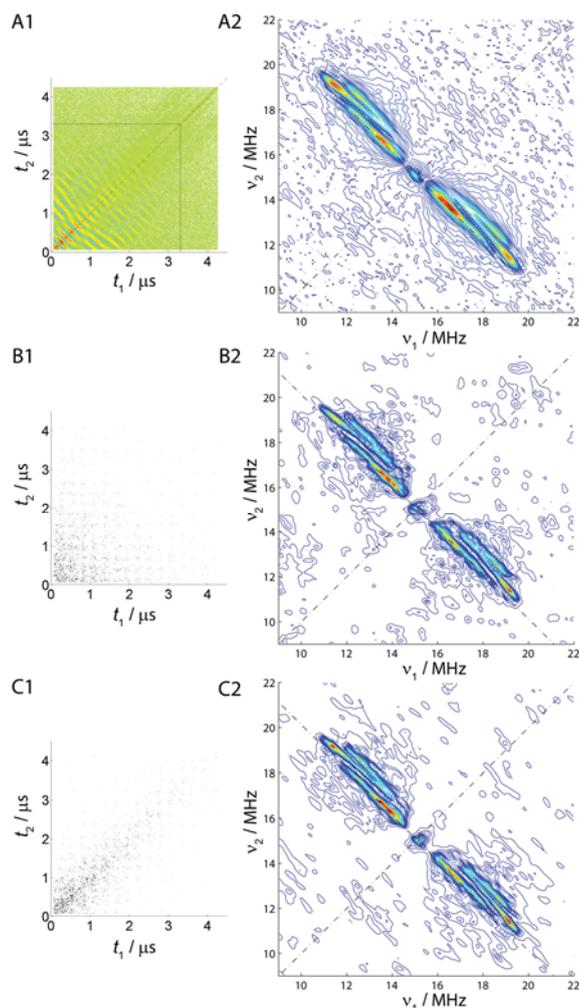


Fig. 2: X-band (9.7822 GHz) HYSCORE data recorded at 15°K from the paramagnetic $[2Fe-2S]^+$ cluster with an $S = 1/2$ spin from palustrisredoxin-B (PuxB).⁸ (A1) full time-domain matrix, 350×350 points sampled linearly with $\Delta T = 12$ ns and (A2) the corresponding DFT spectrum showing the 1H region. The dash lines in A1 shows the length of the time trace required to maintain the maximum spectral resolution in the DFT spectrum. (B1) NUS scheme computed from eq. (3) where 5% of the full time-domain matrix is retained, and (B2) the corresponding absolute value MaxEnt spectrum. (C1) diagonally weighted exponential NUS scheme computed from eq. (4) using 5% of the full linearly sampled (LS) time-domain data matrix, and (C2) the corresponding absolute value MaxEnt spectrum.

Our aim is to determine the minimum number of time-domain data points that are required to faithfully reconstruct the frequency spectra shown in Fig. 2-A2 and 4-A2. To investigate this we will re-sample the time-domain data matrix (Fig. 2-A1 and 4-A1) according to a NUS schedule, and use a MaxEnt algorithm to reconstruct the frequency spectrum.

Our NUS schedule was computed according to an exponentially decaying sample density with sampled points distributed as¹⁰

$$w_i^k = \prod_i^k \exp(-a^k t_{k,i}), \quad (1)$$

where parameters a^k are rate constants in the k^{th} dimension and need to be well chosen. Along a dimension the rank of each time point is

$$N_i = R_i^{1/w_i}, \quad (2)$$

where R is a random number between 0 to 1. For the case of HYSCORE, the a^k parameters are the same in both dimensions, so the NUS matrix N_{cr} is

$$N_{cr} = R_{cr}^{1/\exp(-at_{1,c})\exp(-at_{2,r})} \quad (3)$$

This NUS scheme provides a random and exponentially decaying density of points and it has been extensively tested on NMR data where the parameters a^k are optimally adjusted in relation to the transverse relaxation times T_2^* (usually around $1.5 \times T_2^*$).⁴ For HYSCORE an optimal a value can be adjusted relative to the decay of the electron spin echo envelope modulation (ESEEM) resulting from evolution of the nuclear coherences—the wanted information giving rise to cross-peaks in the HYSCORE experiment (see Fig. 3).

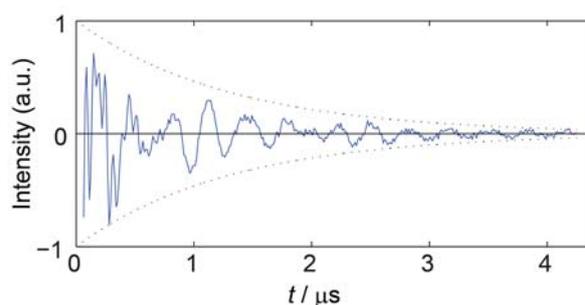


Fig. 3: Decay of the ESEEM trace. Shown is one baseline corrected time trace from the X-band (9.7822 GHz) HYSCORE data recorded from the paramagnetic $[2Fe-2S]^+$ cluster. The dotted lines show the decay of the ESEEM, $1/T_{\text{ESEEM}} \approx 1/1300$ ns.

Once a rank for each point in the HYSCORE time-domain matrix has been computed using eq. (3), we constructed NUS HYSCORE data matrixes at various cut-off levels where e.g. 5% of the full number of experimental data points was retained. These sparse NUS HYSCORE matrixes were then baseline corrected (using the same procedure as for the DFT) and imputed into the MaxEnt algorithm implemented in the Rowland NMR toolkit, with reconstruction parameters chosen as previously described^{11,12} and a frequency domain spectrum computed. The minimum number of NUS points required for the reconstruction was determined by comparing the MaxEnt spectra to the DFT spectrum from the LS data and looking for spectral artefacts and missing peaks.

Depicted in Fig. 2-B1 for the $[2Fe-2S]^+$ cluster is data employing a NUS scheme that retains only 5% (6125 points) of the LS 350×350 point data matrix, and Fig. 2-B2 shows the corresponding 1H region of the frequency domain spectrum computed using the MaxEnt algorithm. A comparison of the spectrum B2 with A2 in Fig. 2 shows that the same information is contained in both spectra, but with peaks in the MaxEnt spectrum being slightly sharper. This is a known feature of the MaxEnt reconstruction algorithm¹³ which scales the signals non-linearly such that weaker signal components are scaled more towards zero than large components, thus suppressing the tails of broad peaks. This also provides a bias towards false negatives (missing peaks) rather than false positives (added peaks), i.e. weak but real peaks are more likely to be indistinguishable from noise rather than false peaks added.

The decay rate of the sampling density depicted Fig. 2-B1 was found to be optimal, and was computed with a decay rate consistent with the decay rate of the ESEEM of the HYSORE time traces, $1/T_{\text{ESEEM}} \approx 1/(1300 \text{ ns})$, see Fig. 3. Reducing the NUS points to less than 5% of the full HYSORE matrix starts to deteriorate the quality of the reconstructed spectrum. Figure S1-3 shows additional results for 10% and 20% sampling levels and a selection of decay rates a in eq. 3.

To provide a meaningful quantification of the time saving using NUS as compared to the conventional linear sampling scheme, the minimum required number of points of the LS data matrix (Fig. 2-A1) needs to be defined. This was achieved by removing end points from the matrix, calculating the DFT spectrum, and examining peak linewidths. This process showed that at least 60% of the data is required to maintain the original resolution, or a LS 271×271 matrix (73,441 points). This region is depicted in Fig. 2-A1 with the dashed lines which intersects the t_1/t_2 axis at $3.3 \mu\text{s}$. Comparing the minimum number of LS points to the minimum number of NUS points shows that the experimental time is reduced to $6125 / 73,441 = 8.3\%$ of the linear sampling scheme. In terms of instrument time, the 271×271 point experiment took 80 minutes to complete, and the NUS time would be 6.6 minutes.

In addition to the NUS of eq. 3, we investigated a NUS scheme emphasising the importance of signal along the diagonal ($t_1 = t_2$) which is where the nuclear coherence echo encoding the cross-peak information is strongest (refocused) in the weak coupling regime[†].

$$N_{cr} = R_{cr}^{1/\exp(at_{1,c})\exp(at_{2,r})\exp(a'(t_{1,c} - t_{2,r}))} \quad (4)$$

This NUS scheme is shown in Fig. 2-C1 and the MaxEnt spectrum in Fig. 2-C2, and as for the NUS scheme of eq. 3, around 5% of the 350×350 data matrix was needed, demonstrating results are not critically dependent upon the choice of these two sampling schedules. Fig. 4 shows a second HYSORE example from $[\text{Rh}(\text{trop2N})(\text{bipy})]\text{OTf}^+$, and it exhibits a very different distribution of peaks in the frequency domain as compared to the previous example, and with a lower SN ~ 40 . In Fig. 4-A1 the LS time-domain matrix comprising 175×175 points is displayed, and Fig. 4-A2 the corresponding DFT spectrum. In this case the frequency resolution was deteriorated if the LS data was truncated to 150×150 points; 150^2 points are thus required to retain peak linewidths that are not lifetime broadened.

Fig. 4-B1 shows an optimal NUS scheme calculated from eq. 3 and employing 10% of the original 175×175 data points with a sampling density rate matched to the decay of the ESEEM, $1/T_{\text{ESEEM}} \approx 1/(700 \text{ ns})$. Fig. 4-B2 shows the corresponding MaxEnt frequency spectrum. This NUS scheme and MaxEnt reconstruction faithfully reproduces the positions of all peaks in the uniformly sampled DFT spectrum, but exhibits narrower linewidths due to the non-linear scaling of the MaxEnt algorithm. Reducing the number of NUS points below 10% started to deteriorate the spectrum quality and consequently in this example we require *ca.* 14% (3063 points) of the data points of an optimized LS matrix (22500 points) to faithfully represent the frequency spectrum. In terms of instrument time, the 150×150 point experiment took 60 minutes to complete, and the NUS time would be 8.2 minutes.

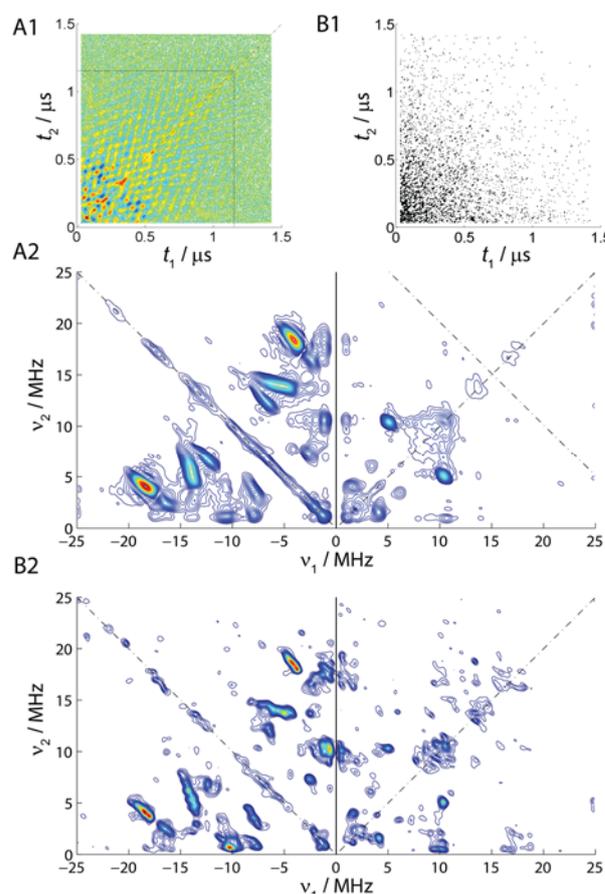


Fig. 4: Q-band (35.309 GHz) HYSORE data recorded at 25°K from the complex $[\text{Rh}(\text{trop2N})(\text{bipy})]\text{OTf}^+$, spin $S = 1/2$.⁹ (A1) full time-domain matrix, 175×175 points sampled linearly with $\Delta T = 8 \text{ ns}$ and (A2) the corresponding DFT spectrum showing ^{14}N signals from three nuclei. The dash lines in A1 shows the length of the time traces required to maintain the maximum spectral resolution in the DFT spectrum. (B1) NUS scheme computed from eq. (3) using 10% of the full LS time-domain data, and (B2) the corresponding absolute value MaxEnt spectrum.

Conclusions

NUS in combination with spectral reconstruction algorithms such as MaxEnt is an accepted and developed method in multi-dimensional NMR spectroscopy. We have shown that this computation approach can be adapted to pulse EPR data for the case of HYSORE spectroscopy, and that the experimental time can be shortened to $<15\%$ of the time required using a conventional linear sampling scheme. For the analysis we employed NUS utilising a pseudo-random exponentially decaying sample density with the optimal decay rate determined relative to the decay rate of the ESEEM. The ESEEM decay rate can be quickly and conveniently measured from 1 trace of a HYSORE experiment before the NUS HYSORE schedule is defined. We investigated a second NUS scheme where points along the diagonal are emphasised since this is where the nuclear coherence echo is refocused for nuclei weakly coupled to the electron spin. Both NUS schemes yielded similar results.

The reconstruction of the frequency domain spectrum from the NUS time domain data was achieved with the Maximum entropy (MaxEnt) algorithms using the implementation in the Rowland NMR toolkit^{11,12}.

The maximum entropy algorithm makes no assumptions about the nature of the signals, and is thus well suited to HYSCORE data where peak shapes cannot be assumed to be e.g. Lorentzian as is required in methods such as linear prediction. Regardless of the reconstruction method employed, the limit of any reconstruction algorithm is to distinguish weak peaks from the noise while ensuring that noise is not identified as real peaks. This is a particularly important consideration for HYSCORE data, where peak intensities (characterised by the modulation depths) are a complicated function of the hyperfine, nuclear quadrupole and nuclear Zeeman interactions¹⁴ and as such signal intensities from different nuclei coupled to the electron spin vary significantly. For this data type, initial results presented here demonstrate that the MaxEnt algorithm is able to handle this intensity variation and accurately reproduce peaks without introducing spectral artefacts. MaxEnt uses entropy as a regularizer to yield a smooth spectrum which is consistent with the experimental data.¹³ The smoothing is non-linear with weaker peaks being scaled down more than stronger peaks. As applied to HYSCORE data this produces peaks and ridges having narrower linewidths as compared to the DFT spectrum. The non-linear property of MaxEnt is in many respects advantageous as peaks in the computed frequency spectrum can be trusted; weak real peaks are more likely to be buried in the noise rather than noise appearing as false peaks (an additional feature of HYSCORE helping to identify real peaks is that cross-peaks symmetric to the diagonal should appear in the spectrum, i.e. $(\nu_{\alpha}, \nu_{\beta})$ and $(\nu_{\beta}, \nu_{\alpha})$).

It has been shown that the reconstruction quality depends critically on using a NUS employing random sampling and to a much lesser extent on the reconstruction methods. In addition to MaxEnt, other methods applicable to NUS HYSCORE data include forward MaxEnt,¹⁵ compressed sensing,¹⁶ iterative thresholding,^{17,18} nonuniform DFT¹⁹ and multi-dimensional decomposition.²⁰

As shown here, the use of NUS sampling enables the HYSCORE experimental time to be significantly shortened, by around an order of magnitude as compared to a linear sampling scheme. This is a very important reduction, as many B_0 field values, defining sets of orientations of the paramagnetic centre with respect to the B_0 , are typically required and at each B_0 setting HYSCORE data with several τ values should be collected to ensure peaks are not suppressed (HYSCORE sequences to suppress the tau dependence have been developed²¹ but have the disadvantage of being less sensitive). As a result a comprehensive HYSCORE study is typically constrained by measurement time. Take the $[2\text{Fe-2S}]^+$ study for example, the EPR spectrum is 450 MHz wide at X-band,⁸ and ideally around 10 HYSCORE B_0 positions should be chosen (a typical $t_{\text{w}/2} = 16$ ns pulse has a 38 MHz excitation profile (f.w.h.h.)), with 3 τ -values per B_0 position, a total of 30 experiments taking ca. 40 hours to complete. Using NUS this time would be reduced to ca. 3.3 hours.

In cases where the sample is weak (e.g. metalloenzyme study), measuring just a single HYSCORE is problematic and can take 12-24 hours. In these cases NUS could be used to increase signal averaging by a factor of 10, amounting to a SN increase of 3.2. The time savings using NUS further enables the possibility to extend HYSCORE experiments into a 3rd dimension by varying τ , thus providing a dimension which correlates peaks to their hyperfine coupling.¹⁴

This work was supported by the Australian Research Council (MM-FT110100925, JRH-FT120100421) and the UQ Summer Student Scholarship scheme (KK).

Notes and references

^a Centre for Advanced Imaging, University of Queensland, St Lucia, QLD 4072. * Address correspondence to Mehdi Mobli (m.mobli@uq.edu.au) and Jeffrey Harmer (jeffrey.harmer@cai.uq.edu.au)

[†] the weak coupling regime is defined as $|A| < 2|\nu|$, where A is the hyperfine coupling and ν the nuclear Zeeman frequency.

Electronic Supplementary Information (ESI) available: Figure S1-S3 and Experimental details. See DOI: 10.1039/c000000x/

1. P. Hofer, A. Grupp, H. Nebenfuhr and M. Mehring, *Chem. Phys. Lett.*, 1986, **132**, 279-282.
2. S. K. Misra, P. P. Borbat and J. H. Freed, *Appl. Magn. Reson.*, 2009, **36**, 237-258.
3. J. A. Kubat, J. J. Chou and D. Rovnyak, *J. Magn. Reson.*, 2007, **186**, 201-211.
4. M. Mobli and J. C. Hoch, *Concepts Magn. Reson. Part A*, 2008, **32A**, 436-448.
5. M. Mobli, M. W. Maciejewski, A. D. Schuyler, A. S. Stern and J. C. Hoch, *Phys. Chem. Chem. Phys.*, 2012, **14**, 10835-10843.
6. J. C. Hoch, M. W. Maciejewski, M. Mobli, A. D. Schuyler and A. S. Stern, *Acc. Chem. Res.*, 2014, **47**, 708-717.
7. J. C. Hoch and A. S. Stern, in *Nuclear Magnetic Resonance of Biological Macromolecules, Pt A*, eds. T. L. James, V. Dotsch and U. Schmitz, 2001, vol. 338, pp. 159-178.
8. J. A. B. Abdalla, A. M. Bowen, S. G. Bell, L. L. Wong, C. R. Timmel and J. Harmer, *Phys. Chem. Chem. Phys.*, 2012, **14**, 6526-6537.
9. T. Buttner, J. Geier, G. Frison, J. Harmer, C. Calle, A. Schweiger, H. Schonberg and H. Grutzmacher, *Science*, 2005, **307**, 235-238.
10. M. Mobli, A. S. Stern, W. Bermel, G. F. King and J. C. Hoch, *J. Magn. Reson.*, 2010, **204**, 160-164.
11. M. Mobli, M. W. Maciejewski, M. R. Gryk and J. C. Hoch, *J. Biomol. NMR*, 2007, **39**, 133-139.
12. M. Mobli, M. W. Maciejewski, M. R. Gryk and J. C. Hoch, *Nat. Methods*, 2007, **4**, 467-468.
13. P. Schmieder, A. S. Stern, G. Wagner and J. C. Hoch, *J. Magn. Reson.*, 1997, **125**, 332-339.
14. A. Schweiger and G. Jeschke, *Principles of Pulse Electron Paramagnetic Resonance*, Oxford University Press, Inc., New York, 2001.
15. S. G. Hyberts, G. J. Heffron, N. G. Tarragona, K. Solanky, K. A. Edmonds, H. Luithardt, J. Fejzo, M. Chorev, H. Aktas, K. Colson, K. H. Falchuk, J. A. Halperin and G. Wagner, *J. Am. Chem. Soc.*, 2007, **129**, 5108-5116.
16. M. Lustig, D. Donoho and J. M. Pauly, *Magn. Reson. Med.*, 2007, **58**, 1182-1195.
17. A. S. Stern, D. L. Donoho and J. C. Hoch, *J. Magn. Reson.*, 2007, **188**, 295-300.
18. S. G. Hyberts, A. G. Milbradt, A. B. Wagner, H. Arthanari and G. Wagner, *J. Biomol. NMR*, 2012, **52**, 315-327.
19. K. Kazimierzczuk, W. Kozminski and I. Zhukov, *J. Magn. Reson.*, 2006, **179**, 323-328.
20. V. Y. Orekhov, I. V. Ibraghimov and M. Billeter, *J. Biomol. NMR*, 2001, **20**, 49-60.
21. L. Liesum and A. Schweiger, *J. Chem. Phys.*, 2001, **114**, 9478-9488.