

Improving resolution in multidimensional NMR using random quadrature detection with compressed sensing reconstruction

M. J. Bostock¹ · D. J. Holland² · D. Nietlispach¹

Received: 31 August 2016 / Accepted: 14 September 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract NMR spectroscopy is central to atomic resolution studies in biology and chemistry. Key to this approach are multidimensional experiments. Obtaining such experiments with sufficient resolution, however, is a slow process, in part since each time increment in every indirect dimension needs to be recorded twice, in quadrature. We introduce a modified compressed sensing (CS) algorithm enabling reconstruction of data acquired with random acquisition of quadrature components in gradient-selection NMR. We name this approach random quadrature detection (RQD). Gradient-selection experiments are essential to the success of modern NMR and with RQD, a 50 % reduction in the number of data points per indirect dimension is possible, by only acquiring one quadrature component per time point. Using our algorithm (CS_{RQD}), high quality reconstructions are achieved. RQD is modular and combined with non-uniform sampling we show that this provides increased flexibility in designing sampling schedules leading to improved resolution with increasing benefits as dimensionality of experiments increases, with particular advantages for 4- and higher dimensional experiments.

Keywords Compressed sensing · Non-uniform sampling · l_1 -norm minimisation · NMR spectroscopy · Random quadrature detection (RQD) · Gradient selection · CS_{RQD}

Introduction

Multidimensional (n D) NMR experiments are indispensable for high resolution NMR spectroscopy studies of macromolecules in biology and chemistry. However, obtaining adequate resolution requires lengthy data sampling that may compromise the achievable sensitivity and lead to extended data collection times.

An area of intense interest for fast NMR spectroscopy involves non-uniform sampling (NUS) of the time domains enabling reduction of the number of acquired time points in all indirect dimensions (Barna et al. 1987; Mobli and Hoch 2014). NUS may be used to improve sensitivity and resolution of NMR experiments compared to their fully sampled equivalents, however the Fast Fourier Transform (FFT) cannot be used to reconstruct the frequency domain spectrum (Palmer et al. 2015). A multitude of different reconstruction methods is available (Orekhov et al. 2001; Kupče and Freeman 2003; Atreya and Szyperski 2004; Tugarinov et al. 2005; Marion 2005; Kazimierczuk et al. 2006; Coggins and Zhou 2008; Matsuki et al. 2009), and recently compressed sensing based techniques (CS) have become popular (Kazimierczuk and Orekhov 2011; Holland et al. 2011; Hyberts et al. 2012).

Nevertheless, despite the improvements introduced by NUS approaches, the $n - 1$ indirect time dimensions of an n D NMR experiment still need to be recorded in quadrature in order to generate high resolution spectra with signals sign-discriminated in frequency and absorptive in line-shape (Keeler and Neuhaus 1985; Ernst et al. 1990).

Electronic supplementary material The online version of this article (doi:10.1007/s10858-016-0062-9) contains supplementary material, which is available to authorized users.

✉ D. Nietlispach
dn206@cam.ac.uk

¹ Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Old Addenbrooke's Site, Cambridge CB2 1GA, UK

² Chemical and Process Engineering Department, University of Canterbury, Christchurch, New Zealand

Quadrature detection is very costly, requiring recording of two data points per indirect time increment, increasing the data collection time by a factor of 2^{n-1} and further compromising the achievable spectral resolution. Maciejewski et al. (2011) suggested random acquisition of phase components (random phase detection (RPD)) with Maximum Entropy (MaxEnt) reconstruction as a sampling reduction strategy for amplitude modulated data, using a partial-component sampling scheme (Schuyler et al. 2015). Although in theory partial-component sampling (recording less than 2^{n-1} quadrature components) is applicable to any NMR experiment, in practice, due to the lack of a suitable reconstruction method, this approach is not available to the majority of modern n D NMR spectroscopy experiments, which typically use gradient-enhanced P- and N-type coherence order selection (so called gradient-selection or phase modulation) (see Theory section). Gradient-selection experiments are prevalent in NMR due to their superior artifact suppression and efficient reduction of large unwanted signals. Amongst the crucial experiments inaccessible to the RPD methodology is the [$^1\text{H}, ^{15}\text{N}$]-TROSY class (Pervushin et al. 1997; Salzmann et al. 1998), which is instrumental for the study of large biomacromolecules.

We introduce a new CS-based algorithm (CS_{RQD}) using a modified version of our in-house developed CS reconstruction method (Bostock et al. 2012), which enables reconstruction of data recorded with a partial-component sampling schedule using either amplitude or phase modulation and name this data reduction strategy random quadrature detection (RQD). Reconstruction of RQD data with CS_{RQD} is applicable to the full range of multidimensional NMR experiments, including those with gradient-enhanced coherence order selection and removes the need for complete quadrature detection in such experiments. The number of data points required is then reduced by a factor of two for every indirect time domain, which is achieved by acquiring only one quadrature component per time increment, with the detected component selected at random. Biomolecular NMR experiments are often limited by sensitivity and therefore require longer recording times; compared to full sampling, RQD enables sampling of the indirect dimensions with superior spectral resolution without the need to increase recording times.

Many NMR experiments are typically already recorded with NUS in order to improve resolution and/or sensitivity. The RQD approach is modular and can be combined with traditional NUS sampling. We show that the combination of RQD and NUS allows increased time-point sampling for a given sampling fraction compared to traditional full-component NUS, which may provide increased resolution

and improved reconstruction properties; the benefits of RQD scale with dimensionality.

Consequently, RQD represents a key recording strategy suitable for all types of multidimensional NMR experiments with the potential to accelerate the sampling or improve resolution and reconstruction properties of every available indirect time domain.

Theory

Compressed sensing reconstruction of NUS data

Compressed sensing (CS) reconstructions have recently become popular in NMR spectroscopy for accurate and rapid reconstruction of NUS datasets using either convex ℓ_1 -norm minimization e.g. iterative thresholding (IT) (Kazimierczuk and Orekhov 2011; Holland et al. 2011; Hyberts et al. 2012) or non-convex approaches using $\ell_{p \rightarrow 0}$ minimisation (Kazimierczuk and Orekhov 2011).

Compressed sensing theory (Logan 1965; Candès et al. 2006; Donoho 2006) describes an approach for solving the system of linear equations

$$\mathbf{Ax} = \mathbf{b} \quad (1)$$

where $\mathbf{A}$ is an $M \times N$ matrix and $\mathbf{x}$ is a vector of length N to be recovered from a set of measurements, $\mathbf{b}$, with $M < N$. For NMR spectroscopy, $\mathbf{A}$ is the inverse Fourier transform at the points sampled, $\mathbf{b}$ is the time-domain data and $\mathbf{x}$ is the frequency domain spectrum. In this case, the equations are underdetermined and (1) has infinitely many solutions. The optimal solution can be obtained by finding the sparsest solution which is consistent with the measured data, i.e. minimising the ℓ_0 -norm, a pseudo-norm defined by:

$$\|\mathbf{x}\|_0 = \sum_i |x_i|^0 \quad (2)$$

where $0^0 = 1$ (Donoho 2006). However, this is computationally intractable (Natarajan 1995). Compressed sensing theory shows that minimising the ℓ_1 -norm, which can be carried out using standard linear processing, can achieve the same result provided the solution is sufficiently sparse:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ subject to } \mathbf{Ax} = \mathbf{b} \quad (3)$$

where

$$\|\mathbf{x}\|_1 = \sum_i |x_i| \quad (4)$$

Non-convex minimisations solve an ℓ_p -norm with $p > 0$ where p is reduced with successive iterations:

$$\|\mathbf{x}\|_p = \left(\sum_i |x_i|^p \right)^{1/p} \quad (5)$$

For data containing noise, or which is compressible but not sparse, the constraint in (3) is relaxed for example to:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ subject to } \|\mathbf{Ax} - \mathbf{b}\|_2 \leq \delta \quad (6)$$

where δ is an estimate of the noise level.

Compressed sensing requires data to be sparse in some basis e.g. the frequency domain for NMR spectra, and to have incoherent sampling with respect to that basis, achieved by selection of an appropriate sampling schedule.

NMR experiments can be successfully undersampled in the indirect time-domains using non-uniform sampling (Barna et al. 1987; Schmieder et al. 1994; Rovnyak et al. 2004), which may be represented as follows for a one-dimensional vector (Maciejewski et al. 2011):

$$z_j = x_j + iy_j, \quad \text{for } j = 0, \dots, N-1$$

$$z_j^{\text{NUS}} = \begin{cases} 0 & \text{if } p_j = 0 \\ x_j + iy_j & \text{if } p_j = 1 \end{cases} \quad (7)$$

where j represents the time increments and $\mathbf{p}$ is the sampling vector i.e. $p_j = 1$ represents sampled points. Strictly speaking, for $p_j = 0$, the point is skipped, i.e. no data is acquired. Thus $\mathbf{z}$ has length given by $\sum_j p_j$.

Compressed sensing reconstruction of RQD data

As well as the requirement to sample to long time points to achieve high resolution in multiple indirect dimensions, NMR spectra also require frequency discrimination and signals with a pure, absorptive, phase. This is achieved using quadrature detection, acquiring two data points per time increment. The total number of points acquired is

$$2^{n-1} \times k_1 \times k_2 \times k_3 \times \dots \times k_n \quad (8)$$

where k_n is the number of points in the n^{th} dimension, and 2^{n-1} results from quadrature detection of the $n-1$ indirect dimensions. Quadrature detection may be achieved using amplitude modulated data, phase modulated data or by oversampling by a factor of two in each indirect dimension (the time-proportional phase incrementation method (TPPI) (Marion and Wüthrich 1983)). In each case, quadrature detection requires a factor of two increase in the number of points per indirect dimension.

Random acquisition of quadrature components, which we generalise as random quadrature detection (RQD) for NMR may be represented in a similar manner to full-component NUS data (7) according to hypercomplex notation (Delsuc 1988; Maciejewski et al. 2011). For a two

dimensional experiment, a matrix of hypercomplex points, $\mathbf{z}$, is acquired:

$$z_{j_1, j_2} = x_{j_1, j_2} + i_1 y_{j_1, j_2} + i_2 r_{j_1, j_2} + i_1 i_2 s_{j_1, j_2} \quad (9)$$

where

$$i_1^2 = i_2^2 = -1$$

$$i_1 \cdot i_2 = i_2 \cdot i_1$$

Assuming the directly acquired dimension is fully sampled as is typically the case, RQD sampling is only implemented in the indirect dimensions; for a 2D this is represented as follows:

$$z_j^{\text{RQD}} = \begin{cases} x_{j_1, j_2} + i_1 y_{j_1, j_2} & \text{if } p_{j_1, j_2} = 0 \\ i_2 r_{j_1, j_2} + i_1 i_2 s_{j_1, j_2} & \text{if } p_{j_1, j_2} = 1 \end{cases} \quad (10)$$

Similar to (7) for $p_{j_1, j_2} = 0$, the point is skipped, i.e. no data is acquired.

Amplitude modulated quadrature detection

Using the States (States et al. 1982) or States-TPPI (Marion et al. 1989) protocol the two quadrature components are represented by cosine and sine modulated datasets. In this case, both components generate an absorption mode spectrum, but without sign discrimination. The random phase detection (RPD) approach, demonstrated using MaxEnt reconstruction (Maciejewski et al. 2011) acquires one phase component for each time-point, selecting either the cosine or sine component at random. This approach is equally possible with standard CS reconstruction solving (6) where $\mathbf{b}$ represents cosine/sine type data (see Results).

Phase modulated quadrature detection (gradient-enhanced spectroscopy)

In contrast, phase modulated data obtained from gradient coherence order selection experiments shows frequency encoding either as $\exp(-i\Omega t_n)$, N-type (echo) or $\exp(+i\Omega t_n)$, P-type (anti-echo) coherence, where Ω is the offset frequency and t_n the time-evolution in the n^{th} indirect dimension. Such datasets give rise to frequency discriminated spectra with each sub-type generating peaks with a phase-twist lineshape i.e. a mixture of absorptive and dispersive components, unsuitable for high-resolution NMR work. Acquiring both components and converting them via linear combination to amplitude-modulated data ($S_{\cos}$ and $S_{\sin}$) generates pure absorption spectra (Davis et al. 1992):

$$S_{\cos}(t) = (\exp(i\Omega t) + \exp(-i\Omega t))/2 = (S_P + S_N)/2$$

$$S_{\sin}(t) = (\exp(i\Omega t) - \exp(-i\Omega t))/2i = (S_P - S_N)/2i \quad (11)$$

Random acquisition of either the P-type or N-type component for each time increment is also possible, but this cannot be processed using the standard CS approach due to the intense artifacts generated by the phase-twist lineshape. However, (11) represents a linear transformation of the P-/N-type data. Therefore, this transformation can be built into the compressed sensing reconstruction by introducing an additional matrix, $\mathbf{U}$, which converts data of the form $S_{\cos}, S_{\sin}$ to S_P, S_N at each iteration, ensuring that the reconstructed frequency domain data at each iteration, $\mathbf{x}$, is constrained via the ℓ_2 -norm term to the raw P-/N-type data, $\mathbf{b}_{PN}$. Equation (6) is reformulated to include this function:
$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ subject to } \|\mathbf{U}\mathbf{A}\mathbf{x} - \mathbf{b}_{PN}\|_2 \leq \delta \quad (12)$$

With this formulation the spectrum is only compared with the components of the P-/N-type data that were sampled. We solve (12) using an iterative thresholding (IT) implementation (Bostock et al. 2012). The modified algorithm, CS_{RQD} , is able to reconstruct data with RQD-sampled gradient-selected time domains as purely absorptive, frequency-discriminated, high resolution spectra.

Of course, NMR experiments that include pulse sequence elements that are generally known by the description of ‘sensitivity enhanced’ or ‘preservation of equivalent pathways (PEP)’ (Cavanagh et al. 1991) that result in the transfer of both orthogonal coherence components modulated by the chemical shift during an evolution period are also suited to RQD and can be reconstructed by CS_{RQD} in analogy to the approach described here for P-/N-type RQD data. This applies also to any single transition-to-single transition polarization transfer (ST)₂PT experiments e.g. the $[^1\text{H}, ^{15}\text{N}]$ -TROSY implementations used in this work. Hence, any strict interpretation of the P-/N-type, gradient-selection or phase modulation terminologies employed throughout this contribution should be relaxed to encompass any of the latter experiment types.

Methods

NMR spectroscopy

NMR experiments were recorded on a Bruker Avance AVIII 800 spectrometer operating at a ^1H frequency of 800 MHz, equipped with a 5 mm TXI HCN/z cryoprobe. Data were collected at 298 K on samples that varied in concentration from 0.4 mM for ^{15}N -labeled RalA-GDP, 0.3 mM for U - $[^2\text{H}, ^{15}\text{N}]$ Ala- $[^{13}\text{CH}_3]$ $[^2\text{H}, ^{13}\text{C}, ^{15}\text{N}]$ Ile δ 1- $[^{13}\text{CH}_3]$ Leu, Val- $[pro\text{-}(R), (S)\text{-}^{13}\text{CH}_3, ^{12}\text{CD}_3]$ -pSRII to 0.25 mM for U - $[^2\text{H}, ^{13}\text{C}, ^{15}\text{N}]$ -labeled S195A-human factor IX. Experiments were recorded as gradient-enhanced implementations of 2D $[^1\text{H}, ^{15}\text{N}]$ -BEST TROSY (Lescop

et al. 2010), 3D $[^1\text{H}, ^{15}\text{N}]$ -BEST TROSY HNCACB (Solyom et al. 2013) and 4D HCCH NOESY (Tugarinov et al. 2005). The key acquisition parameters for each of the experiments that generated the spectra shown in the Figures are given in Tables S1–6. For comparative purposes the individual experiments within the 2D, 3D and 4D series were recorded for equal lengths of time.

Time domain sampling

Evolution times in the indirect dimensions were either sampled in full or using NUS. The NUS sampling schemes were generated using ScheduleTool software (Maciejewski et al. n.d.) or custom written scripts and were either exponentially biased, based on estimates of the expected R_2 values for the indirect dimensions ^1H (4D), ^{13}C (3D, 4D) and ^{15}N (2D) or randomly sampled for the constant time ^{15}N (3D) evolution period.

Frequency discrimination

For data sets with fully sampled and full-component NUS sampled indirect time domains, frequency discrimination in each indirect dimension was obtained either in full quadrature for every sampled time point through recording of both components *i.e.* P-type and N-type components in the case of phase modulation and gradient coherence order selection (Davis et al. 1992) or cosine and sine modulated components for amplitude modulated dimensions in States-TPPI fashion (States et al. 1982; Marion and Wüthrich 1983). In the case of random quadrature detection (RQD), for every sampled time point, only one quadrature component for all indirect dimensions was recorded, reducing the size of the data matrix to 1/2 (2D), 1/4 (3D) or 1/8th (4D) of the hypercomplex matrix and enabling a corresponding increase in acquired time points compared to the same total size of the matrix using full-component NUS. The quadrature component that was recorded was selected in a random manner, using in-house written scripts. Control over the quadrature component to be recorded was obtained via the Bruker VCLIST utility in Topspin. Representative RQD sampling schemes for 2D and 3D experiments are shown in Fig. S1.

Data processing

Fully sampled spectra were processed by Fourier transformation using the Azara software package (W. Boucher, unpublished) while the remaining RQD, NUS and RQD-NUS undersampled experiments were reconstructed using a modification of our in-house developed CS reconstruction methods (Bostock et al. 2012), using MATLAB and Python

and based on the iterative thresholding procedure (IT) as described.

2D and 3D reconstructions were carried out on a multi-user server with 48 AMD 6174 cores with 192 GB RAM using the Python multiprocessing module to run reconstructions over multiple cores. 4D reconstructions were carried out on the Cambridge high performance computing Darwin cluster; each node consists of two 2.60 GHz, eight core, Intel Sandy Bridge E5-2670 processors (sixteen cores in total per node) with 64 GB of RAM (4 GB per core). Code was adapted to use the MPI for Python package (mpi4py) with the Open MPI library. Typical processing times are shown in Table S7.

Display of spectra

Contour levels in the Figures were adjusted to enable a direct comparison of peak intensities between the different spectra in a figure taking into account variations in the number of scans.

Results and discussion

Amplitude-modulated data

As proposed by Maciejewski et al. (2011) partial component sampling of quadrature components still allows the achievement of frequency discrimination, providing the quadrature components are sampled at random. Standard CS processing can reconstruct such spectra by constraining the reconstruction to the acquired cosine/sine components ($\|Ax - b\|_2$ term in (6)). Similar to the previously suggested MaxEnt processing, CS reconstruction of such spectra efficiently suppresses artifacts from the RPD sampling and reproduces all the wanted peaks, generating a spectrum with frequency discrimination (Fig. 1).

Phase-modulated data

As described in the theory section, reconstruction of partial component phase-modulated data requires modification of

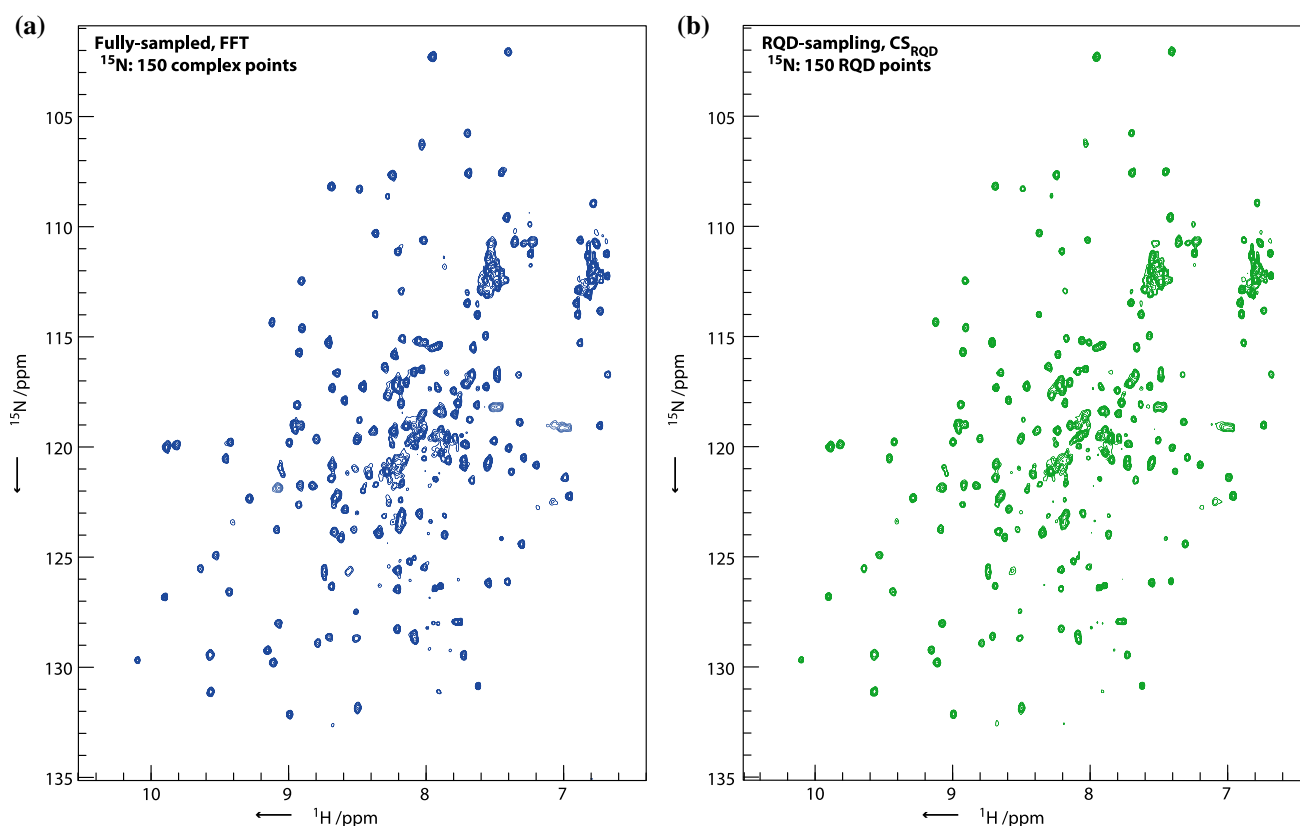


Fig. 1 Reconstruction of a 2D SOFAST [^1H , ^{15}N]-HMQC (Schanda and Brutscher 2005) for RalA-GDP showing the fully-sampled FFT reconstruction of a spectrum recorded with 150 complex points in the ^{15}N dimension (a) in comparison to CS_{RQD} reconstruction of a spectrum recorded with 150 RQD points, i.e. the same $t_{1,\text{max}}$ but selecting either the cosine or sine-modulated component at random for each time-increment, as suggested by Maciejewski et al. (2011)

(b). For comparative purposes, a and b are recorded for the same total experiment time; a is recorded with $n_s = 2$ and b with $n_s = 4$. Due to the different number of scans, spectra are scaled to give the same maximum peak height in a and b. The two experiments are highly similar, with good reproduction of peak positions, shapes and intensities in the RQD spectrum (b). Acquisition parameters are given in Table S1

the standard CS algorithm to ensure the spectrum is constrained to the original P-/N-type data at the acquired data points (CS_{RQD}). CS_{RQD} reconstruction of a 2D gradient-enhanced $[\text{H},^{15}\text{N}]$ -TROSY of the 165 residue G-protein RalA-GDP, acquired with RQD in the ^{15}N dimension, is shown in Fig. 2. This is representative of the high spectral quality obtainable using CS_{RQD} , demonstrating faithful reproduction of peak positions, intensities and line shapes when compared to a conventionally recorded FFT spectrum. Artifact levels in CS_{RQD} reconstructed RQD sampled data sets are generally very low and do not interfere with

any spectral analysis. A substantial benefit of RQD sampling is the ability to increase the spectral resolution in the indirect dimension for a given experiment time (Fig. 2c). For an unbiased comparison all three spectra depicted in Fig. 2 were recorded for the same total amount of time. In the current comparison, RQD sampling enables doubling of the resolution (Fig. 2a–c, inserts). CS_{RQD} reconstruction of RQD sampled data faithfully reproduces peak positions and signal intensities (Fig. 2d, e).

Higher dimensional experiments, e.g. 3D, typically combine gradient-selection in one indirect dimension with

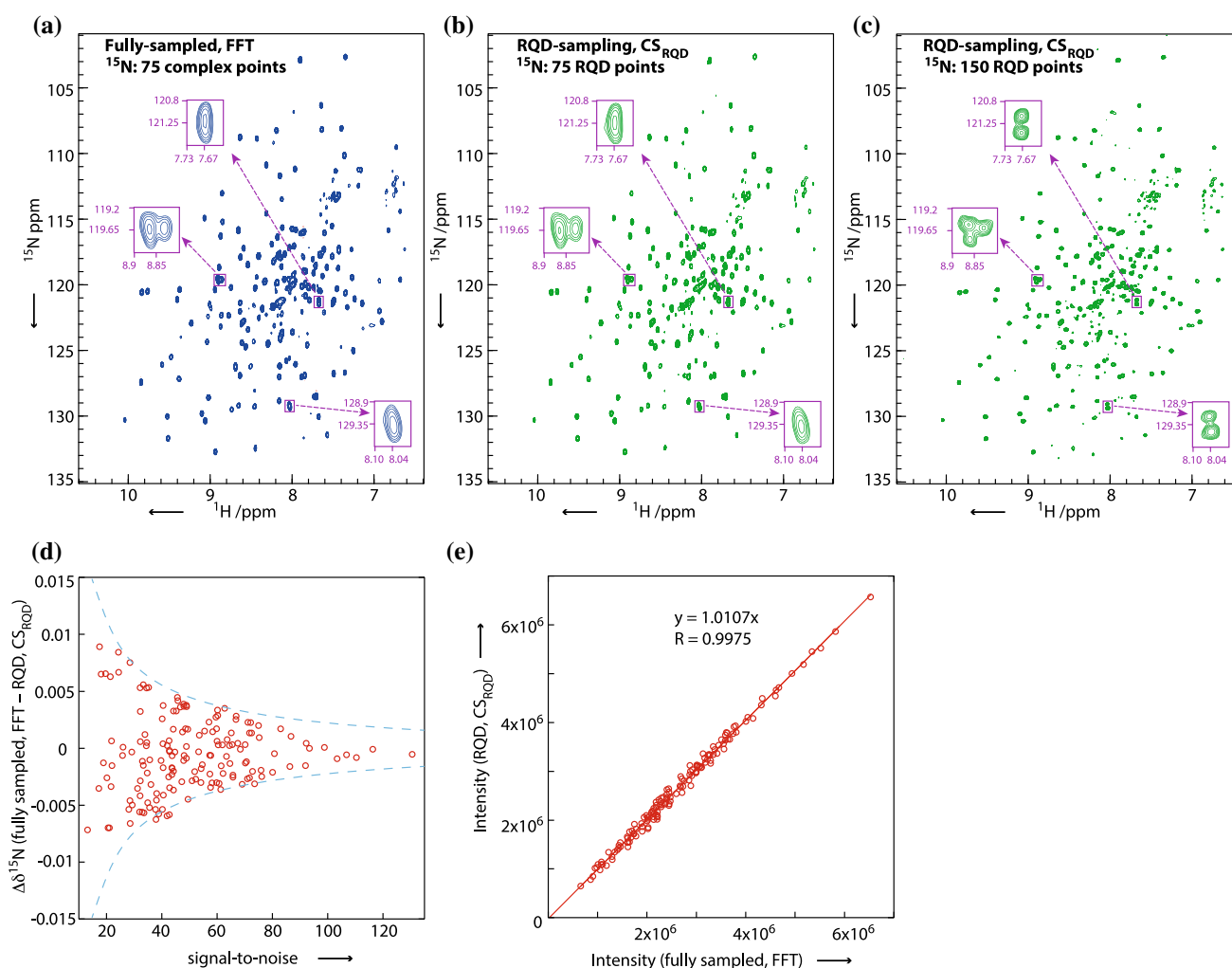


Fig. 2 Comparison of a 2D gradient-enhanced $[\text{H},^{15}\text{N}]$ -TROSY of RalA-GDP showing the fully-sampled FFT reconstruction of an experiment recorded with 75 complex points in the ^{15}N dimension (a) in comparison to the CS_{RQD} reconstruction of a dataset recorded with 75 RQD points, i.e. with the same maximum evolution time but selecting either the P- or N-type component for each time-increment at random (b). In c the time-saving from RQD is used to increase the resolution by recording 150 RQD points. For comparative purposes, a–c are recorded for the same total experiment time; for b this is achieved by doubling the number of scans (Table S2). The purple boxes show three enlarged regions which emphasize the increased

resolution in c obtained through RQD sampling. d, e compare the ^{15}N chemical shift positions and peak intensities from the FFT and CS_{RQD} spectra as red circles. The blue lines in d indicate the ^{15}N chemical shift reproducibility of a fully sampled FFT reconstruction based on line width, acquisition time and signal-to-noise (Kontaxis et al. 2000). The red circles are all well within this reproducibility range. The higher resolution spectra were used for this analysis with 150 complex points ($ns = 4$) for the FFT reconstruction and 150 RQD points ($ns = 8$) for the CS_{RQD} reconstruction, with experiments recorded for equal amounts of time

amplitude modulation in another. RQD can be applied to both indirect dimensions as demonstrated for a 3D [^1H , ^{15}N]-TROSY HNCA (Salzmann et al. 1998) recorded on S195A-human factor IX, a 297 amino acid, 33 kDa protein (Fig. 3) (Johnson et al. 2010). The time saving from RQD allows the resolution to be increased in both indirect dimensions in comparison with the fully sampled FFT experiment.

Partial-component NUS

Although pure RQD may be of use for some higher dimensional ($n \geq 3$) experiments, such experiments are typically already recorded with full-component NUS to reduce data acquisition time and allow improvements in sensitivity and/or resolution. A key question is therefore whether RQD partial-component sampling combined with NUS (RQD-NUS) can outperform standard full-component NUS at a given resolution. This question has been considered theoretically with suggested benefits for partial-component NUS relative to pure NUS due to the increased randomization arising from randomization of the quadrature component in addition to the sampled time points (Schuyler et al. 2015). However, to our knowledge, no comparison in the context of real experimental data has been demonstrated and furthermore, considerations of the partial-component schedules (Schuyler et al. 2015) assume that both components generate an absorptive lineshape, equivalent to applying this approach to RQD-acquired amplitude modulated data (Maciejewski et al. 2011). In reality, this does not account for the challenge of handling the phase-twist lineshape introduced in RQD-acquired phase-modulated data.

Figure 4 shows an NUS sampling schedule for a 3D experiment compared with an RQD-NUS schedule of equivalent resolution. The schedules are displayed in total number of points with the different quadrature components shown in different colours. When considering sampling of time-points and quadrature components in an experiment, for illustrative purposes, these two aspects may be considered separately. In this view full-component NUS is biased towards full-quadrature sampling at the expense of time-point sampling. At the opposite extreme, RQD sampling is biased towards time-point sampling at the expense of sampling the quadrature components. For a full-component NUS schedule, the requirement to sample two components per time point in each indirect dimension reduces the coverage of time points for a given resolution; for an RQD-NUS schedule, two times as many time-points can be sampled per indirect dimension allowing a greater density of coverage. This provides greater flexibility in designing the schedule

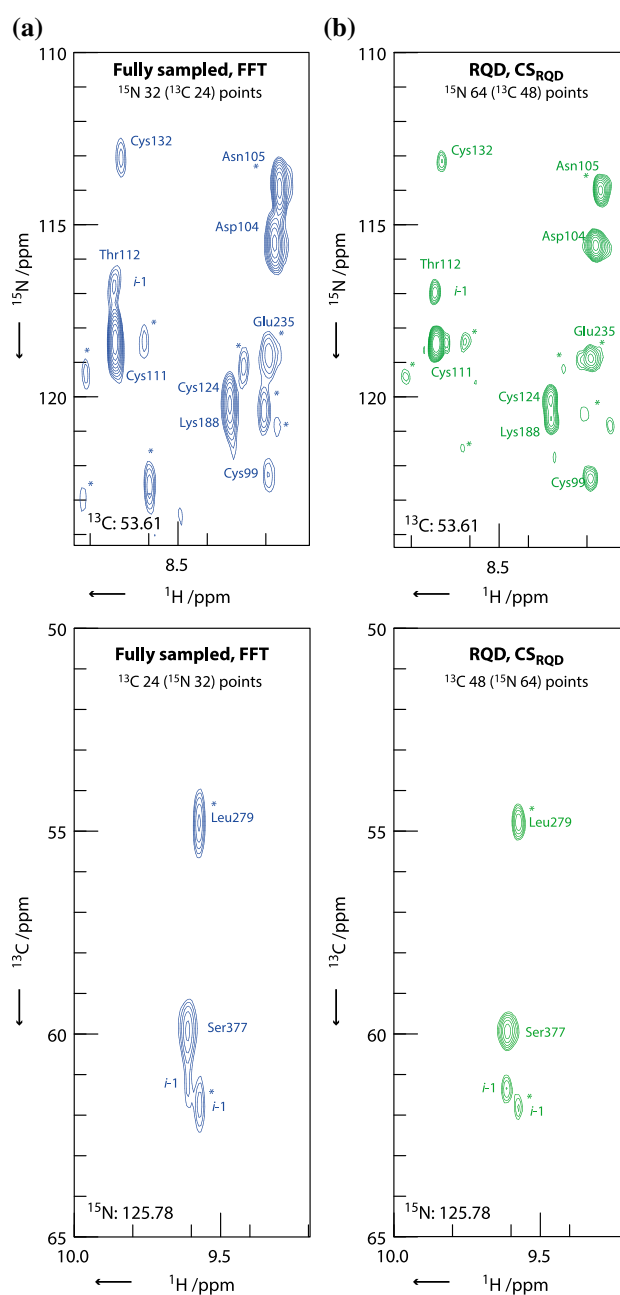


Fig. 3 Selected 2D planes from the reconstruction of a 3D [^1H , ^{15}N]-TROSY HNCA experiment recorded on S195A-human factor IX. **a** 2D [^1H , ^{15}N] and [^1H , ^{13}C] planes from a fully sampled, FFT reconstructed experiment with 32×24 complex points in the ^{15}N and ^{13}C dimensions respectively. **b** The equivalent planes from the CS_{RQD} reconstructed experiment where RQD sampling provides a factor of two saving in each indirect dimension, used here to increase the resolution to 64×48 time-points. Both experiments are recorded for the same total experiment time. Peaks indicated by an asterisk are breakthrough contributions from adjacent planes. The magnitude of the breakthrough peaks in the RQD spectrum is significantly reduced, consistent with the higher resolution of the RQD spectrum. Acquisition parameters are given in Table S3

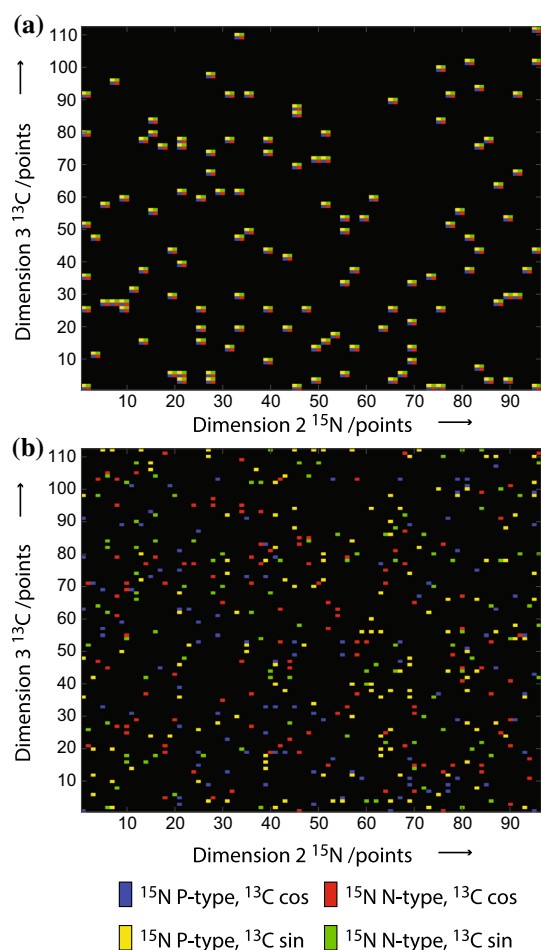


Fig. 4 Comparison of NUS and RQD-NUS sampling schedules for 3D HNCACB data (Fig. 5b). Both schemes acquire the same total number of data points, however the RQD-NUS scheme is biased towards recording more time increments due to the factor of two reduction in quadrature component sampling required in each indirect dimension. Both schemes are drawn from the same exponential sampling distribution function. In **a** and **b**, the different quadrature components are represented with *different colours*, as indicated by the key

and may enable improved resolution due to the greater sampling density at longer time-points compared to an equivalent NUS schedule. Figure 5 compares peaks from a 3D $[^1\text{H}, ^{15}\text{N}]$ TROSY HNCACB experiment recorded either using NUS or RQD-NUS, sampled in both indirect dimensions to equivalent apparent $t_{1,\text{max}}$ in each case (Tables S4 and S5). These examples demonstrate the increased resolving power of the RQD-NUS experiment, allowing peaks overlapped in the NUS spectrum to be distinguished for the same data acquisition time.

For 3D spectra, RQD-NUS allows greater flexibility in point distribution when designing a sampling schedule, in this case resulting in improved resolution for a number of peaks; nevertheless in other parts of the spectrum where the

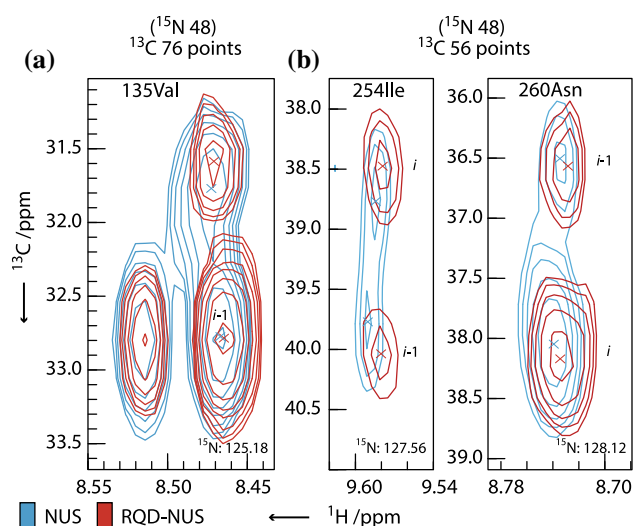


Fig. 5 Selected 2D $[^1\text{H}, ^{13}\text{C}]$ planes from a 3D $[^1\text{H}, ^{15}\text{N}]$ -TROSY HNCACB experiment recorded on S195A-human factor IX using either full-component NUS with CS reconstruction or RQD-NUS sampling with CS_{RQD} reconstruction. **a** a 2D $[^1\text{H}, ^{13}\text{C}]$ plane from a CS-reconstructed experiment with 4.8 % sampling equivalent to a $t_{1,\text{max}}$ of 48×76 complex points in the ^{15}N and ^{13}C dimensions, respectively, recorded with $n_s = 96$. **b** Two 2D $[^1\text{H}, ^{13}\text{C}]$ planes from a CS-reconstructed experiment with 4.5 % sampling equivalent to a $t_{1,\text{max}}$ of 48×56 complex points in the ^{15}N and ^{13}C dimensions, respectively recorded with $n_s = 192$. Both NUS and RQD-NUS experiments are recorded for the same total experiment time. Acquisition parameters are given in Tables S4 and S5

resolution is not limiting there is no visible difference. However, as the dimensionality increases, the density of time-point coverage may need to be reduced substantially for full-component NUS in order to acquire a high resolution experiment in a given recording time. An example of a full-component NUS schedule for a 4D experiment is shown in Fig. 6a. The sampling fraction is 1 %, but since eight quadrature points must be acquired (two per indirect dimension) the reconstruction quality may be poor since so few time-points are characterised. In this situation reducing the bias towards quadrature components with partial-component NUS may provide greater benefits. RQD-NUS provides an eight-fold increase in time-point coverage as seen in Fig. 6b. This may be critical for successful spectral reconstruction at such low sampling density. Examples shown in Fig. 7 for a gradient-enhanced 4D HCCH-NOESY experiment with 1 % sampling compare the CS-NUS reconstruction with CS_{RQD} reconstruction of RQD-NUS data. The full-component NUS reconstruction fails to detect many of the important NOE cross peaks, which are essential for successful structure determination. The sampling distributions used for these reconstructions are shown in Fig. 6; for a fair comparison, the NUS schedule was generated by removing at random 87.5 % of the points

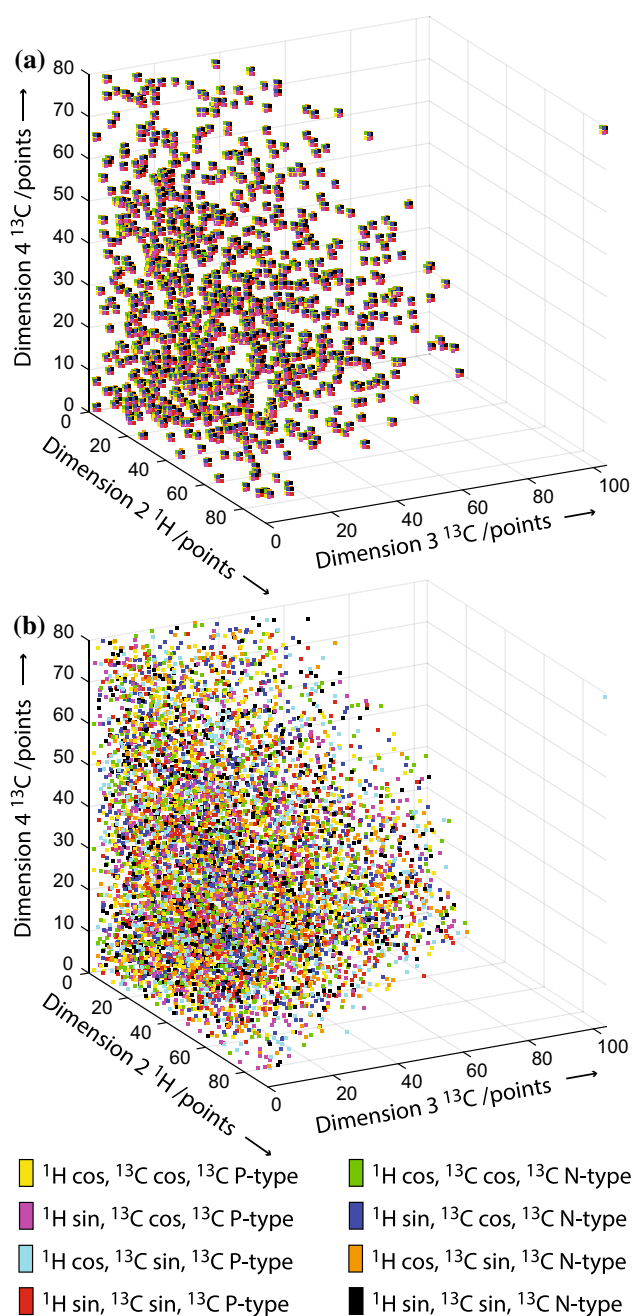


Fig. 6 Comparison of NUS and RQD-NUS sampling schedules for a 4D HCCH (Fig. 7). **a** Full-component NUS and **b** RQD-NUS schedules with 8000 total points (1 % sampling). Both schedules are based on the same exponential sampling distribution function; the NUS schedule was generated by removing 87.5 % of the points from the RQD-NUS schedule and making the remaining 1/8th of the points into full-component quadrature points. The axes show total points in each dimension. The eight quadrature components are shown in different colours

from the RQD schedule and replacing these with full quadrature detection at each remaining time-point. The higher density of time-point sampling in the three indirect dimensions for the RQD schedule resulted in the higher

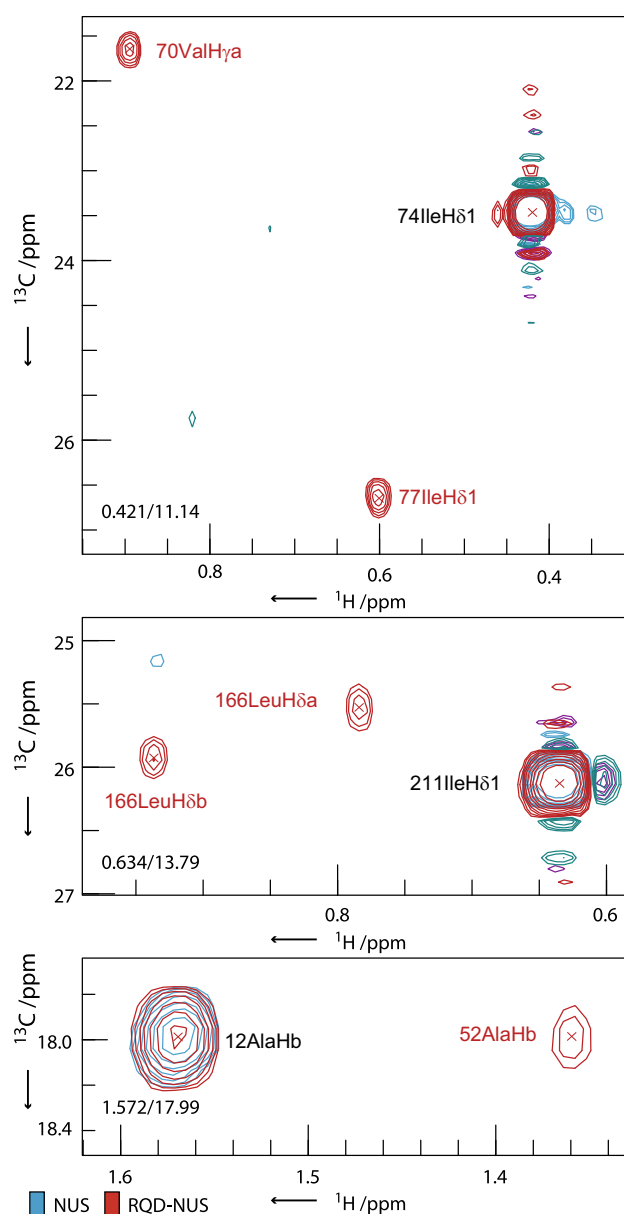


Fig. 7 Selected 2D [^1H , ^{13}C] planes (f_1, f_3) from the reconstruction of a gradient-enhanced 4D HCCH NOESY experiment recorded on ILVA methyl-protonated pSRII. The NUS-only experiment (blue/purple) is recorded with 1000 points from a matrix of $46 \times 52 \times 40$ complex points (1 % sampling). The RQD-NUS (red/green) version was recorded for an equivalent time with 8000 time-points due to the factor of eight undersampling of quadrature components i.e. 1 % overall sampling. The cross peaks are indicated with red text. Full experiment details are given in Table S6. The sampling schedules used are illustrated in Fig. 6

performance of this method. Similar results were also observed using different distributions of NUS points (Fig. S2), indicating that this is not the effect of a single sampling distribution (Fig. S3).

Conclusions

In conclusion, RQD partial-component sampling with CS reconstruction is a powerful method to remove the requirement for full quadrature detection in multidimensional NMR. RQD with CS_{RQD} is applicable to both phase and amplitude modulated data and its benefits are readily available to the full suite of modern NMR experiments. Such experiments are typically gradient-enhanced including the important TROSY-based sequences used for high molecular weight studies. RQD allows a 50 % reduction in the number of data points required per indirect dimension. This can significantly shorten higher dimensional experiments compared to their fully-sampled equivalents allowing the time saved to be converted into substantial resolution enhancements. When compared to full-component CS-NUS reconstructions recorded to equivalent apparent values for $t_{1,max}$ the examples shown here for a 3D experiment demonstrate the potential of RQD to improve peak resolution. As the dimensionality increases, RQD-NUS schedules provide greater coverage of the time points in the $n - 1$ indirect dimensions, which may prove critical for successful spectral reconstruction. We expect further benefits for even higher dimensional experiments. Hence, RQD is of substantial benefit for biomolecular applications, particularly of large proteins or protein complexes, where signal overlap is a key limitation, and higher dimensional experiments are essential to NMR studies. RQD may also be used as a tool for time-saving in situations of high sensitivity e.g. for small molecules where the length of the experiment is determined by the required resolution (sampling limited). However, since RQD sampling diminishes the signal-to-noise ratio (SNR) by a factor of $\sqrt{2}$ for every indirect dimension, shortening an experiment through RQD is only recommended in situations of good SNR. Of course RQD sampling is not limited to acquiring a single quadrature component at each time point; many other sampling scenarios can be envisaged where some time-points have full quadrature detection, others acquire one quadrature component and some time-points are skipped. Analysing the relative benefits of such schedules will be an important topic of future research. Although the experiments used to demonstrate RQD in this paper focus on proteins, the approach is general and will benefit any atomic resolution study that uses multidimensional NMR experiments.

Acknowledgments Thanks to Dr. Arooj Shafiq for the RalA-GDP sample, to Dr. Jennifer Kopanic for the S195A-Factor IX sample, Dr. Duncan Crick for the pSR11 sample and to Dr. Wayne Boucher and Dr. Jenny Barna for assistance with coding. Part of this work was performed using the Darwin Supercomputer of the University of Cambridge High Performance Computing Service (<http://www.hpc.cam.ac.uk/>), provided by Dell Inc. using Strategic Research

Infrastructure Funding from the Higher Education Funding Council for England and funding from the Science and Technology Facilities Council.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Atreya HS, Szyperski T (2004) G-matrix Fourier transform NMR spectroscopy for complete protein resonance assignment. *Proc Natl Acad Sci USA* 101:9642–9647. doi:[10.1073/pnas.0403529101](https://doi.org/10.1073/pnas.0403529101)
- Barna JCJ, Laue ED, Mayger MR, Skilling J, Worral SJP (1987) Exponential sampling, an alternative method for sampling in two-dimensional NMR experiments. *J Magn Reson* 73:69–77. doi:[10.1016/0022-2364\(87\)90225-3](https://doi.org/10.1016/0022-2364(87)90225-3)
- Bostock MJ, Holland DJ, Nietlispach D (2012) Compressed sensing reconstruction of undersampled 3D NOESY spectra: application to large membrane proteins. *J Biomol NMR* 54:15–32. doi:[10.1007/s10858-012-9643-4](https://doi.org/10.1007/s10858-012-9643-4)
- Candès EJ, Romberg J, Tao T (2006) Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans Inf Theory* 52:489–509. doi:[10.1109/TIT.2005.862083](https://doi.org/10.1109/TIT.2005.862083)
- Cavanagh J, Palmer AG, Wright PE, Rance M (1991) Sensitivity improvement in proton-detected two-dimensional heteronuclear relay spectroscopy. *J Magn Reson* 91:429–436. doi:[10.1016/0022-2364\(91\)90209-C](https://doi.org/10.1016/0022-2364(91)90209-C)
- Coggins BE, Zhou P (2008) High resolution 4-D spectroscopy with sparse concentric shell sampling and FFT-CLEAN. *J Biomol NMR* 42:225–239. doi:[10.1007/s10858-008-9275-x](https://doi.org/10.1007/s10858-008-9275-x)
- Davis AL, Keeler J, Laue ED, Moskau D (1992) Experiments for recording pure-absorption heteronuclear correlation spectra using pulsed field gradients. *J Magn Reson* 98:207–216. doi:[10.1016/0022-2364\(92\)90126-R](https://doi.org/10.1016/0022-2364(92)90126-R)
- Delsuc MA (1988) Spectral representation of 2D NMR spectra by hypercomplex numbers. *J Magn Reson* 77:119–124. doi:[10.1016/0022-2364\(88\)90036-4](https://doi.org/10.1016/0022-2364(88)90036-4)
- Donoho DL (2006) Compressed sensing. *IEEE Trans Inf Theory* 52:1289–1306. doi:[10.1109/TIT.2006.871582](https://doi.org/10.1109/TIT.2006.871582)
- Ernst RR, Bodenhausen G, Wokaun A (1990) Principles of Nuclear Magnetic Resonance in one and two dimensions. Oxford University Press, Oxford
- Holland DJ, Bostock MJ, Gladden LF, Nietlispach D (2011) Fast multidimensional NMR spectroscopy using compressed sensing. *Angew Chem Int Ed* 50:6548–6551. doi:[10.1002/anie.201100440](https://doi.org/10.1002/anie.201100440)
- Hyberts SG, Milbradt AG, Wagner AB, Arthanari H, Wagner G (2012) Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling. *J Biomol NMR* 52:315–327. doi:[10.1007/s10858-012-9611-z](https://doi.org/10.1007/s10858-012-9611-z)
- Johnson DJD, Langdown J, Huntington JA (2010) Molecular basis of factor IXa recognition by heparin-activated antithrombin revealed by a 1.7-Å structure of the ternary complex. *Proc Natl Acad Sci USA* 107:645–650. doi:[10.1073/pnas.0910144107](https://doi.org/10.1073/pnas.0910144107)
- Kazimierczuk K, Orekhov VY (2011) Accelerated NMR spectroscopy by using compressed sensing. *Angew Chem Int Ed* 50:5556–5559. doi:[10.1002/anie.201100370](https://doi.org/10.1002/anie.201100370)

- Kazmierczuk K, Koźmiński W, Zhukov I (2006) Two-dimensional Fourier transform of arbitrarily sampled NMR data sets. *J Magn Reson* 179:323–328. doi:[10.1016/j.jmr.2006.02.001](https://doi.org/10.1016/j.jmr.2006.02.001)
- Keeler J, Neuhaus D (1985) Comparison and evaluation of methods for two-dimensional NMR spectra with absorption-mode line-shapes. *J Magn Reson* 63:454–472. doi:[10.1016/0022-2364\(85\)90236-7](https://doi.org/10.1016/0022-2364(85)90236-7)
- Kontaxis G, Clore GM, Bax A (2000) Evaluation of cross-correlation effects and measurement of one-bond couplings in proteins with short transverse relaxation times. *J Magn Reson* 143:184–196. doi:[10.1006/jmre.1999.1979](https://doi.org/10.1006/jmre.1999.1979)
- Kupče Ě, Freeman R (2003) Projection-reconstruction of three-dimensional NMR spectra. *J Am Chem Soc* 125:13958–13959. doi:[10.1021/ja038297z](https://doi.org/10.1021/ja038297z)
- Lescop E, Kern T, Brutscher B (2010) Guidelines for the use of band-selective radiofrequency pulses in hetero-nuclear NMR: example of longitudinal-relaxation-enhanced BEST-type ^1H - ^{15}N correlation experiments. *J Magn Reson* 203:190–198. doi:[10.1016/j.jmr.2009.12.001](https://doi.org/10.1016/j.jmr.2009.12.001)
- Logan BF (1965) Properties of high-pass signals. Columbia University, New York
- Maciejewski MW, Fenwick M, Schuyler AD, Stern AS, Gorbatyuk V, Hoch JC (2011) Random phase detection in multidimensional NMR. *Proc Natl Acad Sci USA* 108:16640–16644. doi:[10.1073/pnas.1103723108](https://doi.org/10.1073/pnas.1103723108)
- Maciejewski MW, Gorbatyuk V, Hoch JC n.d. Sample scheduler for non-uniform data sets. http://sbtools.uchc.edu/nmr/sample_scheduler/. Accessed 25 Aug 2016
- Marion D (2005) Fast acquisition of NMR spectra using Fourier transform of non-equispaced data. *J Biomol NMR* 32:141–150. doi:[10.1007/s10858-005-5977-5](https://doi.org/10.1007/s10858-005-5977-5)
- Marion D, Wüthrich K (1983) Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^1H - ^1H spin-spin coupling constants in proteins. *Biochem Biophys Res Commun* 113:967–974. doi:[10.1016/0006-291X\(83\)91093-8](https://doi.org/10.1016/0006-291X(83)91093-8)
- Marion D, Ikura M, Tschudin R, Bax A (1989) Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *J Magn Reson* 85:393–399. doi:[10.1016/0022-2364\(89\)90152-2](https://doi.org/10.1016/0022-2364(89)90152-2)
- Matsuki Y, Eddy MT, Herzfeld J (2009) Spectroscopy by integration of frequency and time domain information for fast acquisition of high-resolution dark spectra. *J Am Chem Soc* 131:4648–4656. doi:[10.1021/ja807893k](https://doi.org/10.1021/ja807893k)
- Mobli M, Hoch JC (2014) Nonuniform sampling and non-Fourier signal processing methods in multidimensional NMR. *Prog Nucl Magn Reson Spectrosc* 83:21–41. doi:[10.1016/j.pnmrs.2014.09.002](https://doi.org/10.1016/j.pnmrs.2014.09.002)
- Natarajan BK (1995) Sparse approximate solutions to linear systems. *SIAM J Comput* 24:227–234. doi:[10.1137/S0097539792240406](https://doi.org/10.1137/S0097539792240406)
- Orekhov VY, Ibraghimov IV, Billeter M (2001) MUNIN: a new approach to multi-dimensional NMR spectra interpretation. *J Biomol NMR* 20:49–60. doi:[10.1023/A:1011234126930](https://doi.org/10.1023/A:1011234126930)
- Palmer MR, Suiter CL, Henry GE, Rovnyak J, Hoch JC, Polenova T, Rovnyak D (2015) Sensitivity of nonuniform sampling NMR. *J Phys Chem B* 119:6502–6515. doi:[10.1021/jp5126415](https://doi.org/10.1021/jp5126415)
- Pervushin K, Riek R, Wider G, Wüthrich K (1997) Attenuated T_2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci USA* 94:12366–12371
- Rovnyak D, Frueh DP, Sastry M, Sun ZY, Stern AS, Hoch JC, Wagner G (2004) Accelerated acquisition of high resolution triple-resonance spectra using non-uniform sampling and maximum entropy reconstruction. *J Magn Reson* 170:15–21. doi:[10.1016/j.jmr.2004.05.016](https://doi.org/10.1016/j.jmr.2004.05.016)
- Salzmann M, Pervushin K, Wider G, Senn H, Wüthrich K (1998) TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. *Proc Natl Acad Sci USA* 95:13585–13590. doi:[10.1073/pnas.95.23.13585](https://doi.org/10.1073/pnas.95.23.13585)
- Schanda P, Brutscher B (2005) Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *J Am Chem Soc* 127:8014–8015. doi:[10.1021/ja051306e](https://doi.org/10.1021/ja051306e)
- Schmieder P, Stern AS, Wagner G, Hoch JC (1994) Improved resolution in triple-resonance spectra by nonlinear sampling in the constant-time domain. *J Biomol NMR* 4:483–490. doi:[10.1007/BF00156615](https://doi.org/10.1007/BF00156615)
- Schuyler AD, Maciejewski MW, Stern AS, Hoch JC (2015) Nonuniform sampling of hypercomplex multidimensional NMR experiments: dimensionality, quadrature phase and randomization. *J Magn Reson* 254:121–130. doi:[10.1016/j.jmr.2015.02.015](https://doi.org/10.1016/j.jmr.2015.02.015)
- Solyom Z, Schwarten M, Geist L, Konrat R, Willbold D, Brutscher B (2013) BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins. *J Biomol NMR* 55:311–321. doi:[10.1007/s10858-013-9715-0](https://doi.org/10.1007/s10858-013-9715-0)
- States DJ, Haberkorn RA, Ruben DJ (1982) A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants. *J Magn Reson* 48:286–292. doi:[10.1016/0022-2364\(82\)90279-7](https://doi.org/10.1016/0022-2364(82)90279-7)
- Tugarinov V, Kay LE, Ibraghimov I, Orekhov VY (2005) High-resolution four-dimensional ^1H - ^{13}C NOE spectroscopy using methyl-TROSY, sparse data acquisition, and multidimensional decomposition. *J Am Chem Soc* 127:2767–2775. doi:[10.1021/ja044032o](https://doi.org/10.1021/ja044032o)