J Biomol NMR (2010) 47:65–77
DOI 10.1007/s10858-010-9411-2

ARTICLE

Iterative algorithm of discrete Fourier transform for processing randomly sampled NMR data sets

Jan Stanek · Wiktor Koźmiński

Received: 27 October 2009 / Accepted: 24 March 2010 / Published online: 7 April 2010

Abstract Spectra obtained by application of multidimensional Fourier Transformation (MFT) to sparsely sampled nD NMR signals are usually corrupted due to missing data. In the present paper this phenomenon is investigated on simulations and experiments. An effective iterative algorithm for artifact suppression for sparse on-grid NMR data sets is discussed in detail. It includes automated peak recognition based on statistical methods. The results enable one to study NMR spectra of high dynamic range of peak intensities preserving benefits of random sampling, namely the superior resolution in indirectly measured dimensions. Experimental examples include 3D ^{15}N- and ^{13}C-edited NOESY-HSQC spectra of human ubiquitin.

Keywords Multidimensional NMR spectroscopy · Fourier transformation · Sparse sampling · Random sampling · NOESY · Proteins · Ubiquitin

Introduction

Modern NMR techniques allow one to study detailed structure, dynamics and interactions of complex biomolecules such as proteins and DNA in physiological solutions. With constantly increasing size of molecules that scientists are interested in, NMR spectra have become crowded and illegible. The problem could be solved by spreading thousands of resonances into additional spectral dimensions, however, to retain benefits of good signal separation in multidimensional spectra, reasonable resolution ought to be maintained.

The conventional way of recording and processing of multidimensional spectra is based on equispaced sampling of time domain. It therefore requires the fulfillment of Nyquist theorem in each spectral dimension. The theorem stands that for each period of oscillatory signal at least two points should be measured. In other words, the minimum number of *complex* data points N is proportional to spectral width sw and maximum evolution time (Eq. 1).

$$N = sw \times t_{\max} \tag{1}$$

Unfortunately, this results in *exponential* growth of experiment duration with the number of dimensions. Too large sampling intervals lead to observation of aliased peaks. To avoid it one usually decreases maximal evolution time $t_{\max}$ at a cost of desired spectral resolution. As a result, resolution obtainable in 3D spectra, even for relatively fast relaxing proteins, is usually limited. Additionally, constant technical improvement, resulting in higher magnetic field strengths (and thereby wider spectral bands) enhances the significance of sampling limitations. The common concern of slow data acquisition has brought a number of solutions throughout the last two decades.

In general, one can accelerate the measurement using *single-scan* technique which employs spatial encoding of chemical shifts (Frydman et al. 2002). Additionally, in many experiments acquisition rate may be increased by enhancement of recovery of amide protons longitudinal magnetization (Schanda et al. 2006). Different approaches employ *sparse sampling* of the evolution time space, which requires additional efforts at the stage of signal processing. Increasing sensitivity, possible thanks to application of modern equipment like cryoprobes and high-field magnets, is of great importance to those methods. Basically, sparse

J. Stanek · W. Koźmiński (✉)
Faculty of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland
e-mail: kozmin@chem.uw.edu.pl

sampling cannot improve measurements that are sensitivity-limited (Szyperski et al. 2002).

The simplest case of sparse sampling, based on the idea of Accordion Spectroscopy (Bodenhausen and Ernst 1981), is realized by simultaneous evolution of two or more chemical shifts (radial sampling). After Fourier processing the information is retained in the projection spectrum of lower dimensionality that requires additional analysis in order to extract the proper resonance frequencies (Ding and Gronenborn 2002; Kim and Szyperski 2003; Koźmiński and Zhukov 2003; Hiller et al. 2005). On the other hand, a set of radially sampled data can be used to reconstruct the spectrum of full dimensionality. As plain additive back-projection suffers from relatively strong artifacts, more sophisticated methods such as lower-value algorithm (Kupče and Freeman 2003a, b multiway decomposition (Malmodin and Billeter 2005a, b) and iterative frequency identification (Eghbalnia et al. 2005) have been employed. Apart from deterministic methods of spectral reconstruction, statistical and other approaches were proposed with maximum entropy reconstruction (Hoch and Stern 1996), filter diagonalization (Mandelshtam et al. 1998; Armstrong et al. 2005) and multidimensional decomposition (Orekhov et al. 2003; Luan et al. 2005) being most popular. Moreover, the application of polynomial interpolation in conjunction with usual 1D FFT algorithms has been studied in the case of sparse sampling (Marion 2005).

Recently, a general method, referred to as multidimensional Fourier Transformation (MFT), has been proposed for processing multidimensional NMR signals (Kazimierczuk et al. 2006). The estimate of the spectrum *H(f)* is produced by means of evaluating Fourier integrals (which turn into sums in the discrete case), e.g. in 1D case:

$$H(f) \approx \sum_{k=1}^{N} h(t_k) \exp(-2\pi i f t_k) \qquad (2)$$

what corresponds to Monte-Carlo method of numerical integration (Press et al. 2007,§ 7.7). The unique feature of MFT is the ability to process data recorded using any sampling scheme which enables one to search for optimal sampling schedule (Kazimierczuk et al. 2008a). MFT allows to process randomly sampled signals without prior assumptions regarding their bandwidths. In this case, spectral artifacts are present even above Nyquist density; this should be considered a consequence of inherent lack of data rather than improper processing. Despite the presence of spectral artifacts, MFT has been applied for coupling constants measurements (Kazimierczuk et al. 2008b) and 4D experiments for backbone resonance assignment (Zawadzka-Kazimierczuk et al. 2010). These kinds of application do not require additional treatment to reduce artifact level. The simple concept of MFT was also utilized for processing multidimensional NMR spectra by other groups (Coggins and Zhou 2007; Marion 2006).

However, in the case of NOESY spectra, essential for determination of 3D structure of biomolecules, numerous difficulties arise for following reasons:

- the presence of weak signals requiring sufficient sensitivity,
- a high (and a priori unknown) number of resonances causing frequent peak overlaps and ambiguity of cross-peak assignment.

The first issue is rather related to hardware capabilities and sample preparation methods, however, if non-uniform sampling is employed to reach longer evolution times, sensitivity decreases due to relaxation and chemical exchange processes. Second problem can be partially solved by heteronuclear editing and increasing dimensionality of spectrum. This, however, is in contradiction to the desire of high resolution when conventional method of sampling is applied due to sampling limitations. Great efforts have been made to apply non-linear sampling in order to improve spectral resolution of 3D and 4D NOESY spectra using MDD method (Orekhov et al. 2003; Luan et al. 2005; Tugarinov et al. 2005). This created a very attractive alternative to maximum entropy reconstruction (MER), which is applicable rather to acceleration of triple-resonance backbone assignment experiments (Rovnyak et al. 2004) than high quality NOESY spectra due to its inherent nonlinearity and large number of apparent resonances (Orekhov et al. 2003). More recently, a modification of MER was proposed to process more complex data sets such as 3D and 4D NOESY (Hyberts et al. 2009). Apart from the methods mentioned above, an innovative iterative procedure employing information from empty regions in frequency domain (SIFT) has been proposed and applied for processing of 2D HSQC spectrum (Matsuki et al. 2009).

In this communication an alternative approach to processing of multidimensional NMR signal is presented. Our algorithm is derived from CLEAN processing scheme known in radioastronomy (Högbom 1974); examples of its application in NMR have been reported (Barna et al. 1988; Kupče and Freeman 2005; Kazimierczuk et al. 2007; Coggins and Zhou 2008). The new realization of the general CLEAN idea is aimed at automatic and substantial reduction of artifacts, apparent in FT spectra due to *random* and *sparse* sampling of NMR signal. In the following, the detailed description of the algorithm is given. Then the results are presented on experimental data including ^{15}N- and ^{13}C-edited 3D NOESY spectra. In the discussion, we compare our results with those obtained by conventional sampling as well as by MDD and maximum entropy method.

Methods

Signal separation algorithm (SSA)

Overview

The artifact-free spectral reconstruction from incomplete data set is equivalent to missing data problem and requires some prior knowledge of processed signal. Fortunately, multidimensional NMR spectra feature low peak density, i.e. the spectral power is concentrated only in specific areas. This enables one to improve the quality of FT spectra by identifying strong spectral components first. Then, they are separated *in the time domain* from the initial signal in order to reveal weaker resonances. However, the complete and reliable peak identification cannot usually be accomplished on the basis of an initial spectrum, which is distorted by sampling artifacts. Thus the iterative scheme, in which new peaks are gradually found, has been developed (see the flowchart on Fig. 1).

The algorithm, for convenience referred to as Signal Separation Algorithm (SSA), can be summarized as follows ('i' denotes the number of iteration):

1. Start with a residual time-domain signal equal to the recorded FID ($s^{(1)}(t)$).
2. Apply Fourier transform (MFT) to the residual signal, $s^{(i)}(t)$, to obtain an estimate of spectrum, $S^{(i)}(f) = \mathcal{F}(s^{(i)})$.
3. Detect the strongest resonances in $S^{(i)}(f)$ using statistical methods (new peaks are indexed with k from 1 to P_i); if none are found proceed to step 6.
4. For each of detected P_i resonances repeat:

   (a) Find the time-domain representation of the particular resonance, $s_k(t)$, by *inverse* DFT or fitting. Its Fourier transform is supposed to match $\mathcal{F}(s^{(i)})$ in the peak's boundary.

   (b) Subtract it from the previous residual FID signal

   $$s^{(i)}(t) \leftarrow s^{(i)}(t) - s_k(t)$$

   The advantage of updating FID after every peak is processed is that the artifact level falls more quickly and subsequent peaks are less influenced by artifacts. Even better removal of mutual interference of peaks can be obtained by repeating above procedure. In the repeated cycle time domain signal of each peak is added to residual FID and operations (a)(b) are performed again.

   At the end of this step, the residual signal equals to:

   $$s^{(i+1)}(t) = s^{(i)}(t) - \sum_{k=1}^{P_i} s_k(t) \qquad (3)$$

   Having FID of the strongest peaks separated, one can detect the existence of weaker resonances. Go to step 2.
5. The last step includes the MFT of the remaining FID signal: $S_{res} = \mathcal{F}(s_{res})$. Then, the resonances previously extracted from the spectrum are simulated on the full

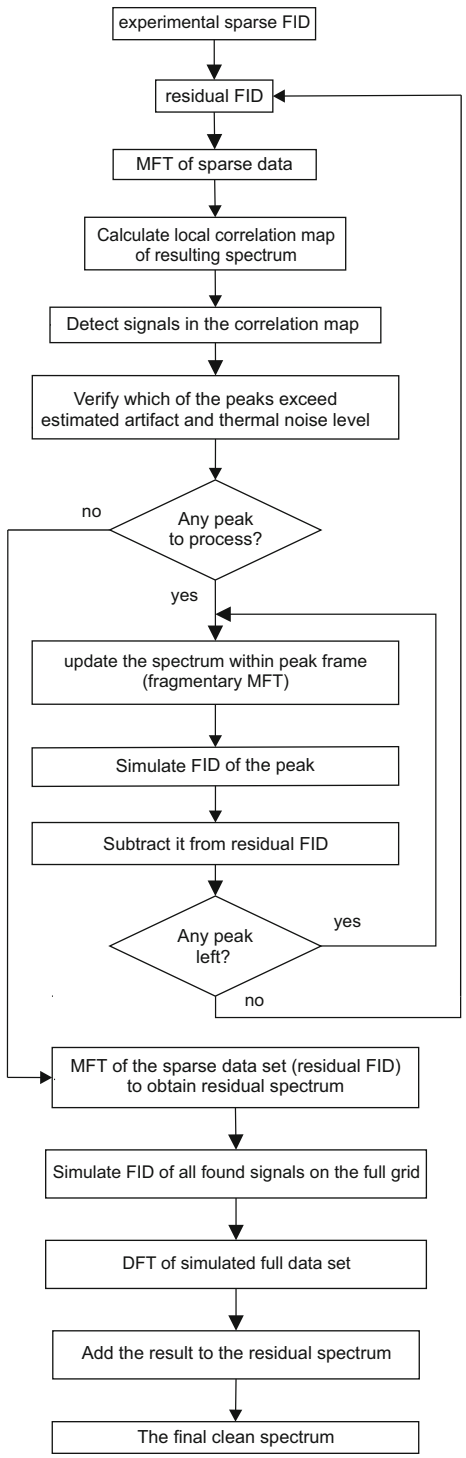

Fig. 1 The flowchart illustrating the principle of iterative signal separation

grid and Fourier transformed. DFT of each peak is scaled to match the intensity observed in the sparse MFT spectrum of that signal ($\mathcal{F}(s_k(t))$) and added to S_{res}.

The final product of the algorithm consists of the strongest peaks and Fourier transform of residual FID that usually contains weak resonances and their (even weaker) artifacts. The great advantage of the design described above is that the processing may be done without the a priori assumption about the number of peaks. On the other hand, it requires proper termination criteria which prevent from detection of false peaks. Further, we discuss the issues associated with the particular stages of algorithm in greater detail.

Signal recognition (Step 3)

The basic strategy for identifying dominant spectral components employs intensity of single spectrum pixels only (Coggins and Zhou 2008). More reliable approach that utilizes a simple statistical method as well as the knowledge of minimal expected linewidths has been developed.

It is widely agreed that randomization of sampling prevents from the occurrence of coherent artifacts. In this case, the quantity defined by Eq. 2 should rather be treated as a statistical spectrum estimator, whose properties have been well described (Kazimierczuk et al. 2009). Therefore, it seems natural to employ statistical methods to distinguish (within some confidence limits) between true peaks and artifacts. As a simple example, a map of local correlation, defined (in the 1D case) as

$$C(f) = \int_{f-\Delta f}^{f+\Delta f} S(f)S(f')df' \rightarrow \sum_{m=i-\Delta}^{i+\Delta} S(f_i)S(f_m) \tag{4}$$

is used in the presented algorithm. The parameter Δf, which determines the spectral width the summation spans over, is evaluated as follows. Prior to the actual processing, the Point Spread Function (Point Response), i.e. the spectrum (its real part) of constant signal of unitary amplitude, is computed:

$$H(f) = \sum_{k=1}^{N} \cos(2\pi f t_k) \tag{5}$$

Then, Δf is estimated as the distance between zero frequency and first zero of the PSF. For example, in the case of uniform random sampling $\Delta f = 1/2\ t_{\max}$ can be analytically found as the first zero of corresponding *sinc* function. To summarize, the purpose of local averaging described above is to utilize the fact that peaks feature a minimal linewidth in FT spectra. It helps to average the noise and determine peak boundaries.

Our simulations showed that correlation maps are extremely selective in terms of signal-to-artifact ratio (S/A). In other words, the cumulative distribution function of pixel intensities (more specifically, their absolute values),

$$p(x) = \frac{\mathrm{Vol}(|C(f)| \leq x)}{\mathrm{Vol}(\mathrm{total})} \tag{6}$$

(where Vol stands for nD volume, or number of pixels in the discrete case, of sufficient intensity) grows very rapidly for low threshold x and reaches plateau (Fig. 2).

Due to obvious nonlinearity of a correlation map, it is relatively easy to derive a fail-safe cutoff level that selects only strong peaks. Our simulations, in which number of indirect points was typical for NMR experiments, allowed to establish the criterion of

$$C_{min} = 50 \times \overline{|C|} \tag{7}$$

where $\overline{|C|}$ stands for the median of absolute values of all pixel intensities in the correlation map (Fig. 3).

Finally, adjacent pixels that exceed the threshold are grouped together to form "peak frames". The latter are rectangular shapes that enclose selected pixels regardless of the particular peak shape or any overlaps. Those *frames* which area is lower than $4\Delta f_1\ \Delta f_2$ (in 2D case) are

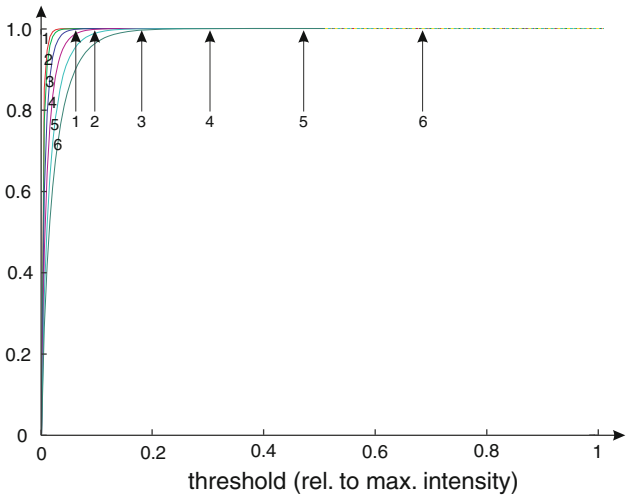

Fig. 2 Cumulative distributions of absolute values of pixels in 2D correlation maps ($p(x)$), obtained in simulations of 4 signals of relative intensities of 1.0 (at −302, 420 Hz), 0.5 (2100, −707 Hz), 0.25 (1502, −1943 Hz), 0.125 (−143, 930 Hz) and equal relaxation rates (75 Hz). White Gaussian noise of σ of 0.0 (curve 1), 0.1 (2), 0.2 (3), 0.3 (4), 0.4 (5), 0.5 (6) was added. The values of threshold determined by the algorithm in each case (1–6) are indicated with *arrows*. Spectral widths of 5 kHz and max. evolution times of 40 ms were set in both dimensions. 1,000 points of uniform random on-grid distribution were used, yielding relative density of 0.025. Sampled signals were processed using plain MFT; 512 complex points were used in each spectral dimension. Cumulative distribution functions were evaluated every 1/200 of corresponding max. correlation value

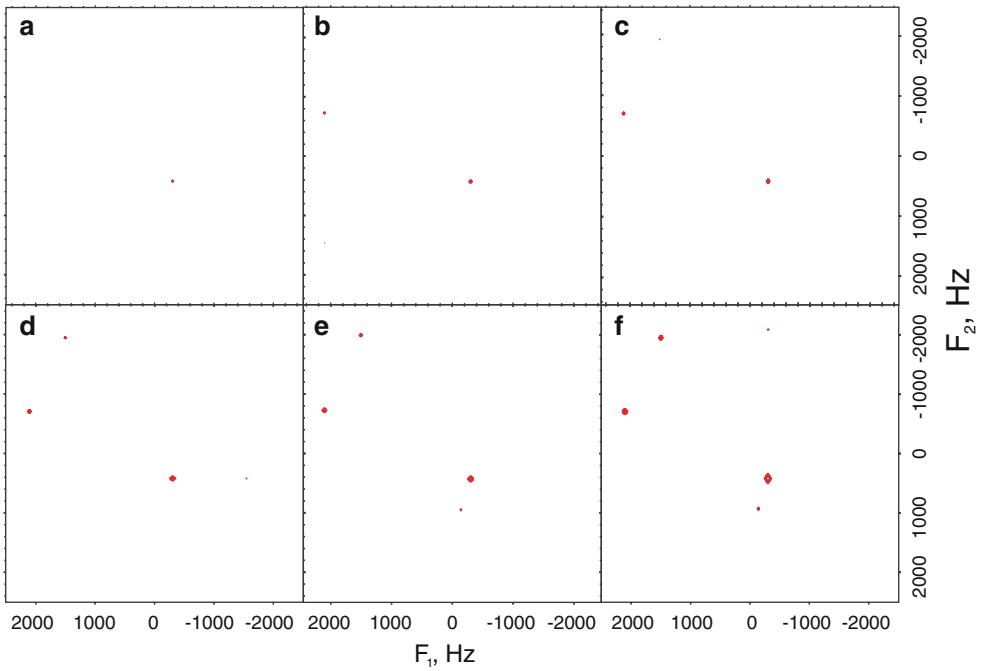

Fig. 3 2D correlation maps obtained in simulations of 4 signals listed in the caption to Fig. 2 (*without* thermal noise), plotted above thresholds determined by Eq. 7. The following number of uniform random on-grid samples were used: 100 (**a**), 400 (**b**), 1000 (**c**), 2000 (**d**), 5000 (**e**), 10,000 (**f**). Spectral widths of 5 kHz and max evolution times of 40 ms were set in both dimensions. Effectively, relative sampling densities were equal to 0.25, 1.0, 2.5, 5.0, 12.5 and 25%, respectively

discarded to ensure stability of further processing steps. It is noteworthy, that described procedure automates the detection of boundaries of strong peaks only. However, it does not indicate weaker peaks that might be heavily obscured by sampling artifacts.

Peak verification (Step 4)

It would be reckless to rely on the signal detection algorithm solely and accept the found peaks without any verification. Therefore, the peak maximum is supposed to fulfill additional criteria:

1. to exceed the threshold of $3 \times \sigma_n$, where σ_n is the standard deviation of thermal noise in the frequency domain. The exact method it is estimated is given below (see "Termination Criteria").
2. to unambiguously exceed the artifact level. The algorithm makes the decision by comparing the peak maximum with the median of absolute values of spectral intensities, $\overline{|I|}$:

$$|I_{peak}| \geq 15 \times \overline{|I|} \tag{8}$$

The median, not average, of pixel intensities is used, as it is less influenced by the presence of intense peaks.

The second threshold can be justified as follows. Firstly, if the peak is considerably affected by either artifacts or thermal noise, its time-domain representation cannot be found with satisfying accuracy. Secondly, if the artifacts or thermal noise are comparable with signal peak, it is pointless to use the presented procedure, as separation of this peak would not significantly decrease the artifact level. Thus, providing that this condition is not fulfilled by any peak, the current iteration should be interrupted. Our experience shows that separation of strong and medium peaks is sufficient to reveal weaker resonances and further efforts are not necessary.

Signal simulation and separation (step 4)

Accurate time-domain signal simulation is essential for our algorithm. Generally, there are two possible approaches:

In the case of single peak one can model the corresponding FID as a monoexponentially decaying oscillatory function $s(t) = A\exp(i\Omega t - R_2 t)$, relying on 3 parameters (or 5 in 2D case). Initial values are set to 0 for relaxation rate R_2 and peak maximum position for frequency offset Ω. Amplitude is adjusted to match the intensity of peak maximum. Then, the parameters are varied to minimize the following function (1D case):

$$w(A, R_2, \Omega) = \sum_{\text{pixels}}^{\text{in frame}} \left(\sum_{k=1}^{N} s(t_k; A, R_2, \Omega) e^{-i\omega t_k} - S^{(i)} \right)^2 \tag{9}$$

In other words, the simulated peak is least-square fitted to obtained spectrum $S^{(i)}$ in the peak boundaries. The single evaluation of $w(A, R_2, \Omega)$ is not computationally demanding, as it includes FID simulation and MFT transformation *to the frame only*. Moreover, the analytic gradient of w is available, allowing for the usage of robust minimization procedures, e.g. Polak-Ribiere (Press et al. 2007, §10.8). It

should be emphasized that the extraordinarily simple model used here is sufficient in most cases and examples of divergence are very rare if peaks are well-separated from each other. There are multiple advantages of the analytic approach:

- The simulated signal is transformed exactly in the same way as experimental data, giving proper lineshape (with regard to applied signal apodization, decaying sampling density and signal truncation). Therefore, this procedure reproduces the artifact pattern accurately. It is impossible to reproduce these factors by simple fitting of assumed shape.
- The peak parameters obtained are modestly affected by thermal noise and artifacts from other peaks as the signal model is rather stiff.

There are situations when the strategy described above fails, namely if the peaks are still overlapped, despite high spectral resolution. In these cases peak fitting procedure might be unstable and more general approach is required.

In principle, an exact time-domain representation of the signal can be obtained by *inverse* DFT

$$h(t_n) = \sum_k \exp(2\pi i f_k t_n) H(f_k) \tag{10}$$

In order to perform inverse FT, imaginary part of the spectrum has to be recovered by application of Hilbert Transform (Bracewell 2000, § 13).

Unfortunately, the contents of frame represents the fragment of MFT spectrum ($S^{(i)}$), not the true one. It is therefore corrupted by sampling artifacts and peak shape is distorted. The interference of artifacts of individual spectral components that compose the frame can be resolved in the following procedure. Let us introduce the following quantities: g—guessed true intensities of peak components, a—the initial contents of the frame (the fragment of $S^{(i)}$) and r—the MFT of the guessed signal:

$$g(f) \overset{iFT}{\Longrightarrow} g(t) \overset{MFT}{\Longrightarrow} r(f)$$

Initially, the following guess is used: $g(f) = a(f)$ (components of $g(f)$ are supposed to have the same sign as the peak maximum; otherwise, they are set to 0). In order to satisfy the condition $r = \mathcal{F}(g(t)) \approx a$ within peak frame the following corrections are applied for each point $'i'$:

$$g_i \leftarrow g_i \times \begin{cases} \frac{2a_i}{a_i + r_i} & \text{if } \operatorname{sgn}(r_i) = \operatorname{sgn}(a_i) \\ 2 & \text{if } \operatorname{sgn}(r_i) \neq \operatorname{sgn}(a_i) \end{cases} \tag{11}$$

The formula can be justified as follows: (1) it amplifies/suppresses the component g_i if the resulting spectrum pixel r_i is weaker/stronger than expected (a_i) or (2) leaves it unchanged otherwise ($r_i = a_i$). However, if the signs of r_i and a_i are different (second case in the formula), the scaling factor cannot be determined and double amplification of g_i is applied until the accordance of sign is obtained. The term a_i in the denominator (and consequently the factor of 2 in the numerator) was added to rectify the possible case of $r_i \approx 0$; otherwise the procedure would cause numerical instability. The iteration is interrupted if (1) the mean square deviation between a and r (per point) is lower than 0.1% of peak maximum or (2) maximum number of 30 iterations is achieved.

The proposed scheme features rapid convergence and stability. For example, in 1977 examples of its use in ^{13}C-edited NOESY (presented below) it required only 8 cycles on average to converge. Further, none examples of divergence (detected as a rise in deviation per point) were detected.

Although the procedure described seems very attractive, it is not preferable as being susceptible for incorporating in g both thermal noise and artifacts from *other* peaks. Its actual use was 10.6 and 25.6% of peaks in the ^{15}N- and ^{13}C-edited NOESY spectra that are presented in this work.

Comparing the second approach to the original CLEAN one can conclude that it removes the need for arbitrary 'loop gain' parameter which may affect linearity and/or computational efficiency or even cause false splittings (Coggins and Zhou 2008). Furthermore, the found signal components, and—more importantly—their artifacts, are completely extracted from the spectrum.

Termination criteria

In the case of iterative algorithms good convergence has to be accompanied by valid termination criteria. These were of the major concern of early CLEAN implementation for processing NMR data (Barna et al. 1988). It was concluded that a fixed limit of iterations is unsuitable approach as it poses a risk of false peaks or cleaning inefficiency. Seemingly, the original CLEAN implementation required a fixed threshold for peak intensities as a stop criterion (Högbom 1974); this was later contradicted as impractical for NMR spectra and replaced by the estimation of remaining "artifact noise" (Coggins and Zhou 2008). Unfortunately, the latter has the disadvantage that the subtraction of false signal could also result in a slight decrease of artifact level and the stopping point is not evident.

Therefore, apart from the strict criteria of peak detection and verification, the estimation of thermal noise was also employed in our algorithm. Firstly, ten indirect t_1t_2 planes (after direct dimension processing) which feature the least mean square intensity (contain no signals) are found, added together and Fourier transformed (MFT). Then, the standard deviation of spectral amplitudes is estimated and divided by $\sqrt{10}$. The resulting value, σ_n, is used as a threshold for peak acceptance as described above. This approach is similar to the estimation of `AIM` parameter in

automatic maximum entropy (MaxEnt) method (Mobli et al. 2007).

Multidimensional Fourier transformation

The mathematical properties of MFT has been discussed thoroughly elsewhere (Kazimierczuk et al. 2009); here the focus is on practical aspects of its use in the presented algorithm.

Sampling restrictions

Contrary to most NUS processing methods, MFT can process data obtained using any sampling scheme; a few variants have been proposed to decrease the artifact level in peak vicinity (Kazimierczuk et al. 2008a). Alas, these concepts are not especially beneficial if numerous peaks are expected. In this case, a true improvement in the sampling procedure should rather be assessed in terms of Signal-to-Artifact ratio (S/A), defined as *the ratio of peak intensity at zero frequency to the mean square amplitude in the outer part of the spectrum of the uniformly sampled constant signal (PSF)* ($sw/2 \geq |f| > sw/4$). Roughly speaking, S/A reflects the relative intensity of artifacts for the given fraction of missing samples. It was found that sampling on Cartesian grid leads to less intense artifacts than without any restrictions imposed (Fig. 4).

For low relative densities θ the difference in S/A is not overwhelming, however, this progress is obtained at no cost. As shown previously, grid restrictions do not lead to the presence of coherent artifacts (Mobli et al. 2006). The only prerequisite for on-grid sampling is the proper estimation of spectral widths, which are fairly well known for proteins. In fact, restriction of signal bandwidth is the actual reason for the increased S/A observed in Fig. 4.

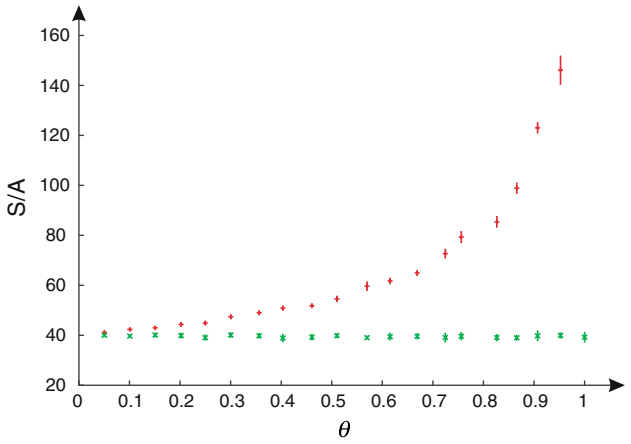

Fig. 4 The plots of spectral Signal-to-Artifact ratio versus relative density, θ, obtained in simulations of randomly sampled constant 2D signal ($\times$ without restrictions, $+$ restricted to Cartesian grid) and processed using MFT. The relative density was changed by varying max. evolution times and keeping spectral widths (both 10 kHz) and number of sampling points (400) constant. Each simulation was repeated 10 times using independent sampling schemes and results averaged out (*error bars* are given). Digital spectral resolution was adjusted in order to eliminate the effect of signal truncation (*sinc shape*), i.e. $\Delta f = 1 / 2t_{max}$. In the case of on-grid sampling, S/A approaches ∞ as $\theta \rightarrow 1$

Performance issues

The use of FFT algorithms to sparsely sampled data (filled with zeros for increments that have not been measured) can greatly speed up the calculation of spectrum. In specific cases, however, this is not the optimal solution and 'slow' FT, i.e. direct evaluation of Fourier sums (Eq. 2) is preferable. This happens if the number of samples (N) is very low in comparison to required resolution (M). 'Slow' FT implemented as matrix-vector multiplication requires exactly MN floating-point operations, whereas FFT accomplishes the task at a cost of $(2M + 1) \log_2 M$ calculations (imaginary parts have to be computed). Therefore, if it reduces the computational burden, one can switch to 'slow' FT at the stage of MFT processing. Another application of 'slow' FT involves a *fragmentary* transform, i.e. the case when only a piece of full spectrum is required, or—equivalently—inverse Fourier transform (Eq. 10) from a peak frame. In both cases (forward or inverse) FFT would have to operate on a full data set thus decreasing the overall performance. To summarize, the effective implementation of computational core of the algorithm that uses both 'slow' and 'fast' FT enables one to develop more complex processing schemes such as SSA algorithm described here.

Results and discussion

Comparison with the previous CLEAN implementation

To show efficiency of artifact suppression and convince that SSA/FT is suitable for the spectra containing peaks of various intensities, a simple one-dimensional simulation has been presented on Fig. 5. This test involves 2 peaks of different amplitude (1:33) and a little amount of white Gaussian noise ($\sigma = 0.01$ relative to the strongest of two peaks). Two methods for separation of signals introduced in this work are compared to the recent CLEAN implementation (Coggins and Zhou 2008) in the regime of low sampling density ($\theta = 0.15$). The latter method was used with default parameters: loop gain equal to 0.3 and noise stability criterion of 5%. Identical criteria of peak acceptance, namely that intensity of a peak needs to be five-fold larger than the standard deviation of remaining noise, were used for all cases. For convenience, noise level was

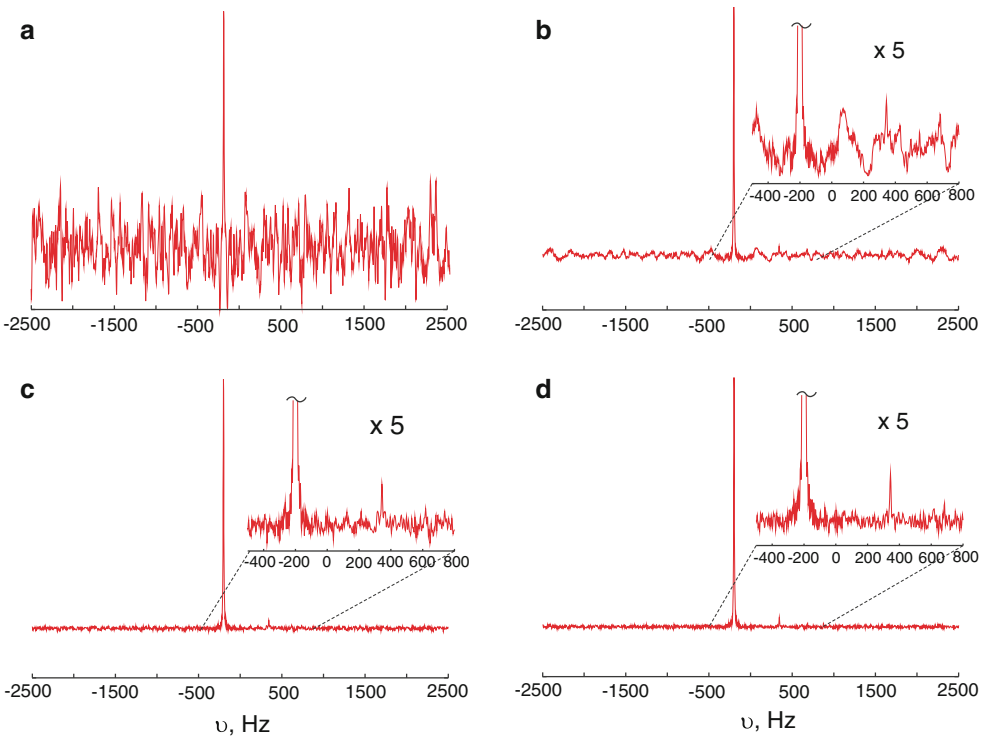

Fig. 5 The comparison of different CLEAN variants: initial spectrum from sparse data (**a**), final results of recent CLEAN implementation (Coggins and Zhou 2008) (**b**) and two methods introduced here, inverse FT with corrections (**c**) and fitting (**d**). Two signals of relative amplitude 1.0 and 0.03, equal relaxation rate of 25 Hz and frequency offset of − 200.3 and 343.23 Hz were simulated. White Gaussian noise of σ = 0.01 was added. 75 points of uniform random on-grid distribution were used. Spectral width of 5000 Hz and max evolution time of 10 ms were set, yielding relative density of ρ = 0.15. 2000 frequency points were used

estimated in the outer part of the spectrum, similarly to S/A calculations described above.

The comparison of initial (Fig. 5a) spectrum to the results of different CLEAN algorithms (Fig. 5b, c, d) convinces of practical effectiveness of this approach. However, the previous implementation seems to perform a little worse than presented here and some amount of artificial noise remains, hampering detection of the weaker resonance. On the other hand, two methods introduced in this work give close results, at least in simple spectra (Fig. 5cd). These results could be explained by the fact that previous CLEAN implementations process the spectrum in

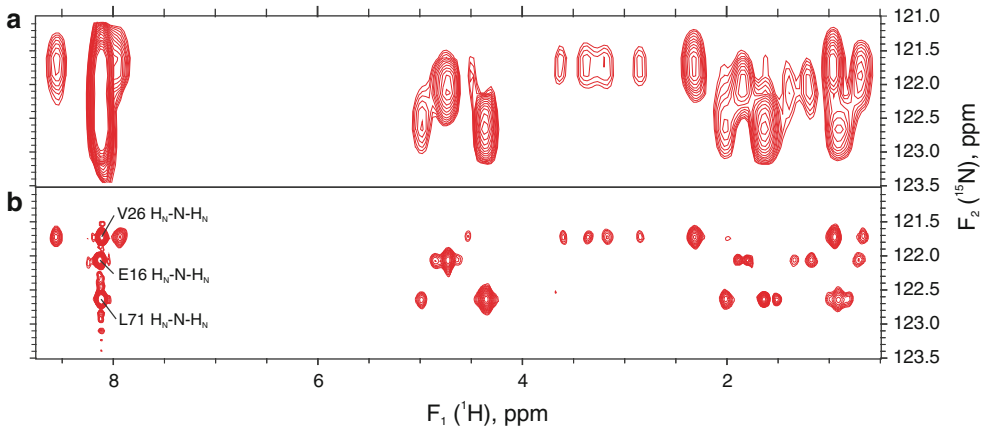

Fig. 6 Contour plots of F_1F_2 cross-sections at F_3 = 8.115 ppm from ^{15}N-edited NOESY-HSQC spectra of ^{13}C,^{15}N-labeled human ubiquitin obtained on 700 MHz spectrometer using conventional (**a**) and sparse on-grid sampling (**b**). Substantial resolution enhancement allowed resolving signals associated with different nitrogen atoms (in E16, V26 and L71 residues). Conventional spectrum was obtained using 82 and 28 (2,296 total) t_1/t_2 increments; spectral widths of 12,000, 2,400 and 12,000 Hz ($F_1 \times F_2 \times F_3$) were assumed, thus maximum evolution times $t_{1,\max}$ = 6.8 ms and $t_{2,\max}$ = 11.7 ms were achieved. Four-fold LP (double in each indirect dimension) was applied. The sparse random time schedule, consisting of 2,300 points, was generated using the *nussampler* from MDD package (with T2 parameters estimated to 0.03 and 0.1 s). In the case of experiment employing random sampling, maximum evolution times of 30 and 100 ms were set in proton and nitrogen dimensions respectively (according to different rates of relaxation), yielding relative density of θ = 0.027. Four scans were coherently added in all four data sets for all data points; as a consequence, acquisition time of both experiments were practically equal. The spectra were transformed with the resolution of 512 × 128 × 2, 048 (**a**) and 2, 048 × 1024 × 2048 (**b**) points in F_1, F_2 and F_3, respectively. The cross-sections are plotted with the same intensity scale

a pixel-by-pixel fashion and are incapable of detect the 'wings' of the resonance within an intensity threshold imposed. The second difference regards the incompleteness of signal extraction due to 'loop gain' parameter different from 1. As a consequence, some spectral power remains at the position of the peak and corresponding artifacts appear in the full spectrum.

Apart from differences discussed above, the new algorithm improves computational effectiveness, as it evaluates full spectrum not after each pixel is extracted but only if the list of P_i found peaks is exhausted. It is quite rare for it to exceed 5 iterations per 2D plane.

Experimental results

The only way to assess the quality of new processing algorithms is to test them on experimental data. The aim of this work was to deal with artifacts in complicated spectra where peaks of various intensities are observed. As a perfect example the heteronuclear-edited NOESY-HSQC experiments are presented, which give crowded spectra even for small proteins. Experiments were performed on a human ubiquitin sample, known for good NMR properties. For this reason, we decided to shorten the acquisition time to a minimum granting thermal sensitivity.

The first overnight experiment (15 h), ^{15}N-edited NOESY HSQC, was carried out in the extreme regime of sampling density ($\theta = 0.027$). This was achieved by setting the maximum evolution times to 30 and 100 ms in proton and nitrogen dimensions, respectively. These tough conditions were deliberately set in order to show excellent resolution and efficiency of artifact suppression. The results are compared to conventional experiment of equal acquisition time.

Conventional way of time domain sampling results in poor resolution especially along indirectly sampled dimensions what might hamper resonance assignment even in small proteins. One of the most severe cases is presented on Fig. 6a. Due to insufficiently resolved F_2 (^{15}N) dimension several peaks overlap and cannot be directly isolated. On the other hand, exploiting the non-uniform sampling in combination with efficient iterative processing enables one to obtain well-resolved spectrum shown in Fig. 6b. It is noteworthy, that all signals present in the Fig. 6a are retained and their chemical shifts can be determined more accurately.

The efficiency of artifact suppression even in complex fragments of ^{15}N-edited NOESY spectrum is demonstrated on Fig. 7. We show that SSA/FT algorithm is capable of reconstruction of virtually all resonances present in the conventional spectrum (Fig. 7a). The results are compared with two other methods that can process non-uniformly sampled data, i.e. maximum entropy reconstruction (Fig. 7b)

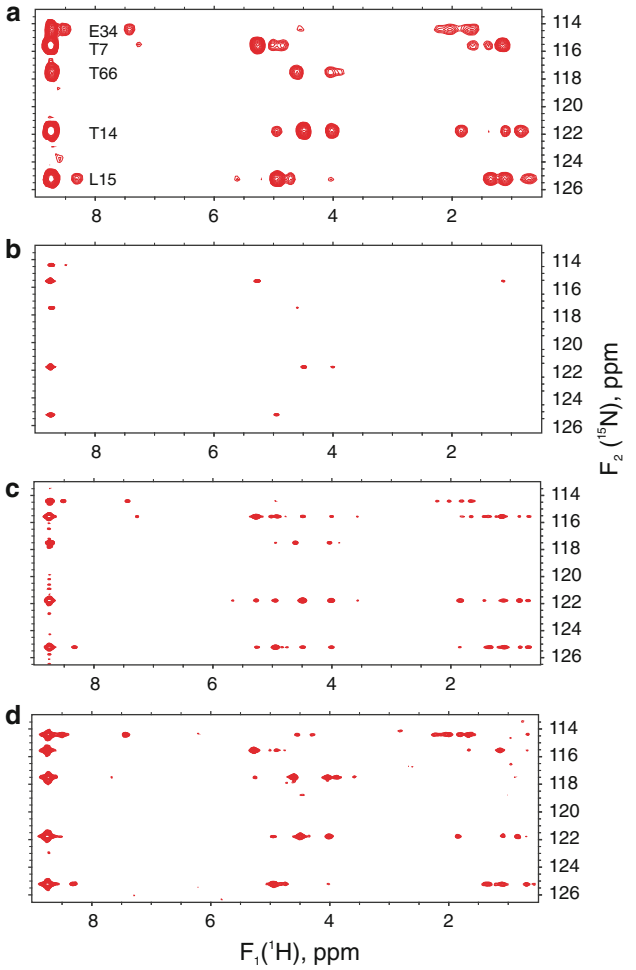

Fig. 7 Comparison of contour plots of F_1F_2 planes at ω_3 (^{1}H) = 8.743 ppm, obtained in ^{15}N-labeled NOESY-HSQC experiment for ^{13}C,^{15}N-labeled human ubiquitin on 700 MHz spectrometer using conventional (**a**) and sparse on-grid sampling (**b–d**). Experimental and sampling details are given in the caption to Fig. 6. Spectra (b-d) were obtained by the analysis of the same sparse data with 3 methods: automatic maximum entropy reconstruction as implemented in *Rowland NMR Toolkit v.3* (**b**); Multidimensional Decomposition, implemented in *mddnmr1.6* (**c**) and SSA/FT algorithm described here (d). Cutoff level of 1.5% relative to the amplitude of well-separated T14 diagonal peak was used for all cases. The spectra were transformed with the resolution of 512 × 128 × 2, 048 (**a**) and 2, 048 × 1024 × 2, 048 (**b–d**) points in F_1, F_2 and F_3, respectively. Square cosine weighting function was used in all dimensions (except for indirect dimensions in MaxEnt reconstruction). Four-fold LP (double in each indirect dimension) was used in the case of conventional spectrum. Conventional spectrum (**a**) is free of artifacts, however it features poor resolution owing to shorter evolution times. On the other hand, the spectra obtained by sparse sampling and different reconstruction methods (**b–d**) do not retain the full information content. This is mainly due to the different relative density requirements; MDD and SSA/FT methods give close results in the conditions described above. Details of MaxEnt and MDD processing are given in the "Materials and methods" section

and Multidimensional Decomposition (Fig. 7c); details of processing are given in section "Materials and methods". It is difficult to draw general conclusions from only one

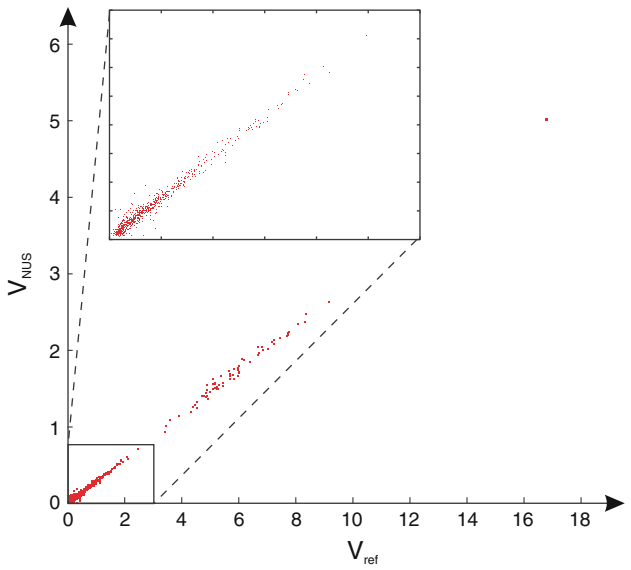

Fig. 8 Correlation plot of peak volumes (in arbitrary units) of diagonal and crosspeaks for 70 out of 76 backbone residues of human ubiquitin (except M1, P19, E24, P37, P38 and G53), measured in randomly sampled (V_{NUS}) and reference ^{15}N-edited NOESY-HSQC spectra of the same resolution (V_{ref}). Peak volumes were calculated with SPARKY utility using Gaussian fit; procedure converged in both spectra for 732 peaks (752 and 769 in reference and randomly sampled ones, respectively). The correlation coefficient $R^2 = 0.998$ was obtained. Volume errors can be explained by the presence of thermal noise and peak overlaps. The reference spectrum was obtained using 240 and 120 t_1/t_2 increments; spectral widths of 12,000, 2,400 and 12,000 Hz ($F_1 \times F_2 \times F_3$) were assumed, thus maximum evolution times $t_{1,\max} = 20$ ms and $t_{2,\max} = 50$ ms were achieved. In the case of randomly sampled experiment, 2,300 on-grid points were randomly chosen using decaying sampling density ($\cos^2(0.5\pi t/t_{\max})$). Equal maximum evolution times as in the reference experiment were set, yielding relative density of $\theta = 0.08$. Four scans were coherently added in four data sets for all data points. The spectra were transformed with the resolution of 2, 048 × 1, 024 × 2, 048 points in F_1, F_2 and F_3, respectively. The cosine-square weighting function was applied in the case of reference experiment

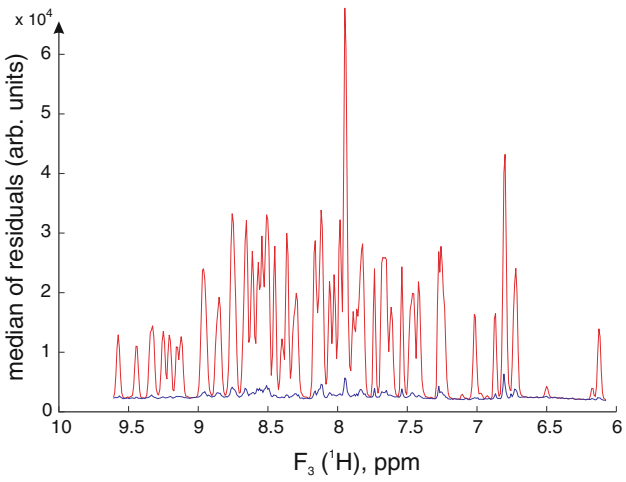

Fig. 9 Noise estimated as median of absolute values of pixels in initial (*the upper curve*) and residual (*the bottom curve*) spectra obtained during cleaning process of randomly sampled ^{15}N-edited NOESY HSQC spectrum. Almost uniform residual noise across the full spectrum (F_3 between 6.1 and 9.6 ppm), achieved in the final iterations of SSA/FT, convinces of excellent artifact suppression. Intensity of thermal noise can be assessed for example in the region between 6.2 and 6.5 ppm. Experimental details are given in the caption to Fig. 6

comparison, however, one can notice that in these particular conditions MaxEnt does not properly reproduce the true spectrum. On the contrary, MDD method, which was originally tested on NOESY spectra obtained with 80–90% time savings, coped quite well with presented demanding data.

The ^{15}N-edited NOESY HSQC experiment was repeated with decreased maximum evolution times ($t_{\max} = 20$ and 50 ms for ^{1}H and ^{15}N, respectively; $\theta = 0.08$). The quality of reconstructed spectrum was investigated by comparison with the reference spectrum, i.e. conventionally sampled one that features the same resolution. Peak picking in the reconstructed spectrum was performed in Sparky using the threshold of 10 × estimated global noise level (σ_{NUS}). The latter was determined in as a median of absolute values of 10^4 randomly chosen spectral intensities. Only peaks that have a corresponding diagonal peak were taken into account. Subsequently, obtained peak list was verified on a basis of reference spectrum. 828 peaks appeared to be properly reconstructed. 11 peaks were classified as false for one of the following reasons: (1) the cross-peak intensity in the reconstructed, I_{NUS}, and the reference spectra, I_{ref} (after appropriate scaling as described below), differed by more than 3 σ_{NUS}; (2) the peak does not appear in the reference spectrum at all.

Additionally, the linearity of the presented method was investigated. The comparison of peak volumes in the reconstructed and reference spectra is shown on Fig. 8. Excellent linear correlation obtained here ($R^2 = 0.998$) shows that sparsely sampled and SSA/FT reconstructed spectra can be used in structural studies of biomolecules.

Finally, an estimation of the number of missing peaks was performed. As the reference spectrum features significantly larger thermal sensitivity (a factor of $\sqrt{1/\theta} \approx 3.5$) the threshold for peak picking was determined relative to peak intensities rather than to noise level. More specifically, the threshold was transferred from reconstructed spectrum (10 σ_{NUS}) using the linear correlation between intensities ($I_{NUS} \propto I_{ref}$ with $R^2 = 0.999$ for all 828 peaks). 14 more peaks were found in the reference spectrum using the same relative threshold.

In order to assess the effectiveness of artifact suppression, one can estimate noise level as a median of absolute value of spectral intensities. The evaluation was performed for each point of directly detected dimension and plotted on the Fig. 9. Significant decrease of noise level in the

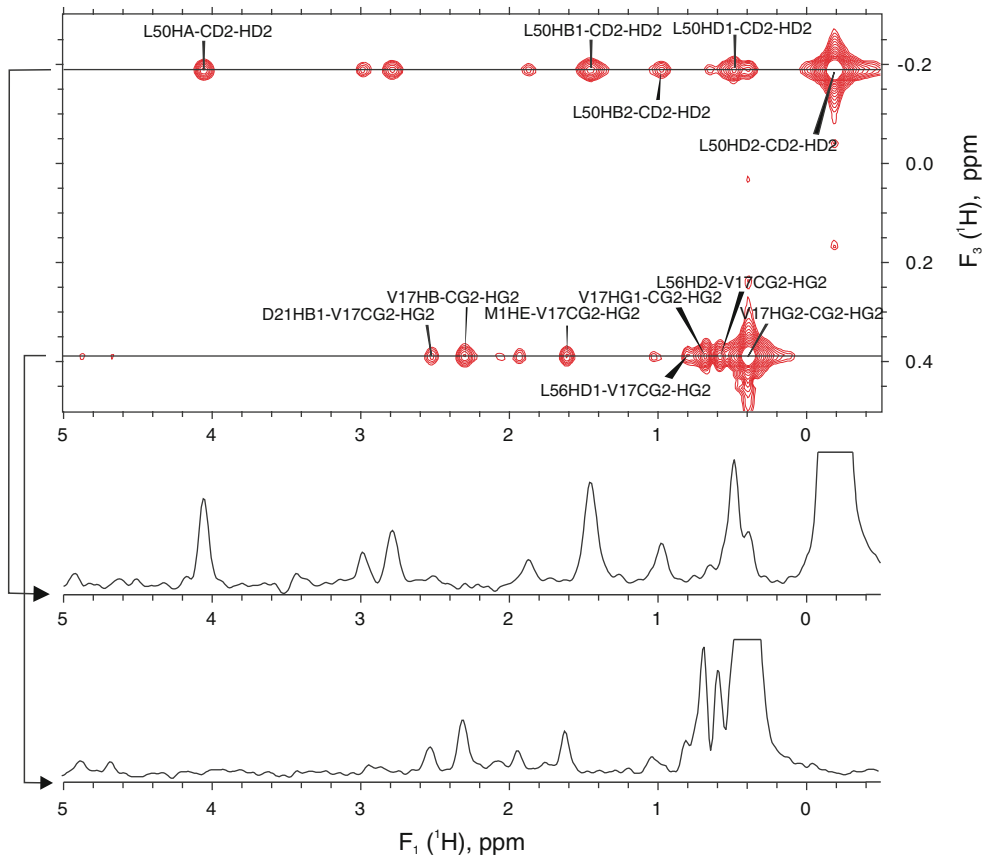

Fig. 10 Contour plot of F_1F_3 plane at $\omega_2(^{13}C) = 19.58$ *ppm*, obtained in ^{13}C-labeled NOESY-HSQC experiment for ^{13}C,^{15}N-labeled human ubiquitin on 700 MHz spectrometer using sparse on-grid sampling and processed by SSA/FT. Cutoff level of 0.5% (relative to L50 HD2 diagonal peak amplitude) was applied. 2700 data points with Gaussian decay of sampling density ($\exp(-t^2/2\sigma^2)$, $\sigma = 0.5$) were used; for each t_1/t_2 point 4 scans were coherently added for 4 hypercomplex data sets. Maximum evolution times of $t_{1,max} = 30$ ms, $t_{2,max} = 12.5$ ms were set. Spectral widths were estimated to 8,000, 14,000 and 12,000 Hz in F_1, F_2 and F_3, respectively, yielding relative density of $\theta = 0.064$. The spectrum was transformed with the resolution of 1,024, 1,024 and 2,048 complex points in F_1, F_2 and F_3, respectively. Unambiguous cross-peaks were assigned on the basis of 5D HC(CC-TOCSY)CONH experiment (Kazimierczuk et al. 2009) and BMRB Entry 5387. 1D crossections are presented to show enhanced resolution in the indirectly detected proton dimension

residual spectrum (i.e. before reintroducing reconstructed strong resonances) in comparison to the corrupted initial one demonstrates high efficiency of the cleaning algorithm.

The 3D ^{13}C-edited NOESY experiment is much more demanding due to larger number of resonances and higher dynamic range. However, such spectra are more important in structural studies. Another difficulty arises due to C-C couplings that make increasing carbon evolution time significantly (and thereby achievable resolution) not especially beneficial ($t_{2,\ max} = 12.5$ ms was used). The experiment duration was set to 18 hours, effectively, the relative density was equal to $\theta = 0.064$. The experiment was carried on ubiquitin sample in H_2O/D_2O (9:1) solution thus overall sensitivity might have decreased. The results of SSA/FT processing presented on Fig. 10 show excellent suppression of artifacts in relation to weak off-diagonal correlation signals. This was possible owing to greater sensitivity of this experiment.

Materials and methods

The 3D ^{15}N- and ^{13}C-edited NOESY-HSQC spectra have been recorded for 1.5 mM ^{13}C, ^{15}N—double labeled human ubiquitin in 9:1 H_2O/D_2O at pH 4.5 at 298 K on a Varian NMR System 700 spectrometer equipped with a Performa XYZ PFG unit and using the 5 mm ^1H, ^{13}C, ^{15}N—triple resonance probehead with high power ^1H, ^{13}C and ^{15}N $\pi/2$ pulses of 5.9, 13.5, and 31.0 μs, respectively. Mixing time of 150 ms was used in all cases. The pulse sequences have been adapted from the Varian Userlib BioPack package.

^{15}N-edited NOESY (presented on Fig. 7) data were processed using automatic maximum entropy (MaxEnt) reconstruction as implemented in the Rowland NMR Toolkit msa2d program (Mobli et al. 2007). Processing script was created by RNMRTK Script Generator v. 3.97 (http://www.sbtools.uchc.edu/nmr/). `DEF` and `AIM`

determination was based on noise in the last FID. Final MSA2D calculation was performed with DEF = 0.055 and constant LAMBDA = 0.84. Resolution was set to 2048 × 1024 × 2048 complex points in F_1, F_2 and F_3 dimensions, respectively.

The same data as above were processed with MDD as implemented in mddnmr v. 1.6. Typical values of parameters lambda = 0.005 and SRSIZE = 0.2 ppm were used. regions.runs script was modified with the information from 2D H/N HSQC to set optimal values of NCOMP parameters in the subsequent subregions. Resolution was set to 2048 × 1024 × 2048 in F_1, F_2 and F_3 dimensions, respectively.

Our processing software was intensively exploiting BLAS (Blackford et al. 2002) and FFTW (Frigo and Johnson 2005) algorithms implemented in Intel Math Kernel Library 10.0 for Linux. All spectra has been saved in the format of SPARKY (Goddart and Kneller 1989–2008) program.

Processing times for were 16 h 37 min, 38 min and 1 h 39 min (using 1 CPU core) on a 64 bit CentOS 5.0 based server equipped with 2 Intel Xeon 1.6 GHz CPUs and 8 GB RAM for MaxEnt, MDD and SSA/FT, respectively.

Conclusions

The main objective of presented approach was to substantially improve spectral resolution using longer evolution times, rather than to shorten acquisition time by omitting most increments. Whereas the latter approach is more appropriate to assignment-oriented experiments, the former one (that we follow) seems more beneficial in anyway time-consuming NOESY.

We have shown that one can successfully process non-uniformly sampled multidimensional NMR signals using Fourier methods. The problem of artificial noise has been tackled by employing iterative procedure of automatic signal identification and separation (SSA). Good results have been achieved in moderate computational time for 3D ^{15}N- and ^{13}C-edited NOESY spectra, featuring high dynamic range of peak intensities.

Enhanced spectral resolution can be of a great importance when dealing with biomolecules more complex than presented here. Further advantages of the method are expected when employed to 4D experiments due to even better signal separation.

Software availability

The program binaries are available from the authors upon request.

Acknowledgments Special thanks are addressed at Maxim Mayzel from Swedish NMR Centre for his help in using MDD package, and at prof. J.C. Hoch from University of Connecticut Health Center for providing access to Rowland NMR Toolkit v.3. This work was supported by grant number: N301 07131/2159, founded by Ministry of Science and Higher Education in years 2006-2009. Research cofinanced by the European Social Fund and State funds under the Integrated Regional Operational Programme, Measure 2.6 "Regional Innovation Strategies and transfer of knowledge", Mazovian Voivodship project "Mazovian Ph.D. Scholarship".

References

Armstrong GS, Mandelshtam VA, Shaka AJ, Bendiak B (2005) Rapid high-resolution four-dimensional NMR spectroscopy using the filter diagonalization method and its advantages for detailed structural elucidation of oligosaccharides. J Magn Reson 173: 160–168

Barna JCJ, Tan SM, Laue ED (1988) Use of CLEAN in conjunction with selective data sampling for 2D NMR experiments. J Magn Reson 78:327–332

Blackford LS, Demmel J, Dongarra J, Duff I, Hammarling S, Henry G, Heroux M, Kaufman L, Lumsdaine A, Petitet A, Pozo R, Remington K, Whaley RC (2002) An updated set of basic linear algebra subprograms (BLAS). ACM Trans Math Soft 28:135–151

Bodenhausen G, Ernst RR (1981) The accordion experiment, a simple approach to three-dimensional NMR spectroscopy. J Magn Reson 45:367–373

Bracewell RN (2000) The Fourier transform and its applications. McGraw-Hill Higher Education, New York

Coggins BE, Zhou P (2007) Sampling of the NMR time domain along concentric rings. J Magn Reson 184:207–221

Coggins BE, Zhou P (2008) High resolution 4-D spectroscopy with sparse concentric shell sampling and FFT-CLEAN. J Biomol NMR 42:225–239

Ding K, Gronenborn AM (2002) Novel 2D triple-resonance NMR experiments for sequential resonance assignments of proteins. J Magn Reson 156:262–268

Eghbalnia HR, Bahrami A, Tonelli M, Hallenga K, Markley JL (2005) High-resolution iterative frequency identification for NMR as a general strategy for multidimensional data collection. J Am Chem Soc 127:12528–12536

Frigo M, Johnson SG (2005) The design and implementation of FFTW3. Proc IEEE 93:216–231

Frydman L, Scherf T, Lupulescu A (2002) The acquisition of multidimensional NMR spectra within a single scan. Proc Natl Acad Sci 99:15662–15858

Goddart TD, Kneller DG (1989–2008) SPARKY 3. University of California, San Francisco

Hiller S, Fiorito F, Wüthrich, Wider G (2005) Automated projection spectroscopy (APSY). P Natl Acad Sci USA 102:10876–10881

Hoch JC, Stern AS (1996) NMR data processing. Wiley, New York

Högbom J (1974) Aperture synthesis with a non-regular distribution of interferometer baselines. Astron Astrophys Suppl 15:417–426

Hyberts SG, Frueh DP, Arthanari H, Wagner G (2009) FM reconstruction of non-uniformly sampled protein NMR data at higher dimensions and optimization by distillation. J Biomol NMR 45:283–294

Kazimierczuk K, Zawadzka A, Koźmiński W, Zhukov I (2006) Random sampling of evolution time space and Fourier transform processing. J Biomol NMR 36:157–168

Kazimierczuk K, Zawadzka A, Koźmiński W, Zhukov I (2007) Lineshapes and artifacts in multidimensional Fourier transform of arbitrary sampled NMR data sets. J Magn Reson 188:344–356

Kazimierczuk K, Zawadzka A, Koźmiński W (2008a) Optimization of random time domain sampling in multidimensional NMR. J Magn Reson 192:123–130

Kazimierczuk K, Zawadzka A, Koźmiński W, Zhukov I (2008b) Determination of spin-spin couplings from ultrahigh resolution 3D NMR spectra obtained by optimized random sampling and multidimensional Fourier transformation. J Am Chem Soc 130:5404–5405

Kazimierczuk K, Zawadzka A, Koźmiński W (2009) Narrow peaks and high dimensionalities: exploiting the advantages of random sampling. J Magn Reson 197:219–228

Kim S, Szyperski T (2003) GFT NMR, a new approach to rapidly obtain precise high-dimensional NMR spectral information. J Am Chem Soc 125:1385–1393

Koźmiński W, Zhukov I (2003) Multiple quadrature detection in reduced dimensionality experiments. J Biomol NMR 26: 157–166

Kupče E, Freeman R (2003a) Projection-reconstruction of three-dimensional NMR spectra. J Am Chem Soc 125:13958–13959

Kupče E, Freeman R (2003b) Reconstruction of the three-dimensional NMR spectrum of a protein from a set of plane projection. J Biomol NMR 27:383–387

Kupče E, Freeman R (2005) Fast multidimensional NMR: radial sampling of evolution space. J Magn Reson 173:317–321

Luan T, Jaravine V, Yee A, Arrowsmith CH, Orekhov VY (2005) Optimization of resolution and sensitivity of 4D NOESY using multidimensional decomposition. J Biomol NMR 33:1–14

Malmodin D, Billeter M (2005a) Multiway decomposition of NMR spectra with coupled evolution periods. J Am Chem Soc 127:13486–13487

Malmodin D, Billeter M (2005b) Signal identification in NMR spectra with coupled evolution periods. J Magn Reson 176:47–53

Mandelshtam VA, Taylor HS, Shaka AJ (1998) Application of the filter diagonalization method to one- and two-dimensional NMR spectra. J Magn Reson 133:304–312

Marion D (2005) Fast acquisition of NMR spectra using fourier transform of non-equispaced data. J Biomol NMR 32:141–150

Marion D (2006) Processing of ND NMR spectra sampled in polar coordinates: a simple fourier transform instead of reconstruction. J Biomol NMR 36:45–54

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

Mobli M, Stern A, Hoch JC (2006) Spectral reconstruction methods in fast NMR: reduced dimensionality, random sampling and maximum entropy. J Magn Reson 182:96–105

Mobli M, Maciejewski MW, Gryk MR, Hoch JC (2007) Automatic maximum entropy spectral reconstruction in NMR. J Biomol NMR 39:133–139

Orekhov VY, Ibraghimov I, Billeter M (2003) Optimizing resolution in multidimensional NMR by three-way decomposition. J Biomol NMR 27:165–173

Press WH, Flannery BP, Teukolsky SA, Vetterling WT (2007) Numerical recipes in C, 3rd edn. Cambridge University Press, Cambridge

Rovnyak D, Frueh DP, Sastry M, Sun ZYJ, 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

Schanda P, Melckebeke HV, Brutscher B (2006) Speeding up three-dimensional protein NMR experiments to a few minutes. J Am Chem Soc 128:9042–9043

Szyperski T, Yeh DC, Sukumaran DK, Moseley HNB, Montelione GT (2002) Reduced-dimensionality NMR spectroscopy for high-throughput protein resonance assignment. P Natl Acad Sci USA 99:8009–8014

Tugarinov V, Kay LE, Ibraghimov I, Orekhov VY (2005) Highresolution four-dimensional H-1-C-13 NOE spectroscopy using methyl-TROSY, sparse data acquisition, and multidimensional decomposition. J Am Chem Soc 127:2767–2775

Zawadzka-Kazimierczuk A, Kazimierczuk K, Koźmiński W (2010) A set of 4D NMR experiments of enhanced resolution for easy resonance assignment in proteins. J Magn Reson 202:109–116