UNIVERSITY OF GOTHENBURG

This is an author produced version of a paper published in **Topics in current chemistry**

This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

GUP

Gothenburg University Publications

http://gup.ub.gu.se

`

Generalized Fourier transform for non-uniform sampled data

Krzysztof Kazimierczuk[a,b], Maria Misiak[a], Jan Stanek[a], Anna Zawadzka-Kazimierczuk[a] and Wiktor Koźmiński[a,*]

[a] Faculty of Chemistry, University of Warsaw, Pasteura 1, 02093 Warsaw, Poland

[b] Swedish NMR Centre, University of Gothenburg , Box 465, S-405 30 Gothenburg, Sweden

[*] To whom correspondence should be addressed:

kozmin@chem.uw.edu.pl

Abstract

Fourier transform can be effectively used for processing of sparsely sampled multidimensional data sets. It provides possibility to acquire NMR spectra of ultra-high dimensionality and/or resolution which allow easy resonance assignment and precise determination of spectral parameters e.g. coupling constants. In this chapter, the development and applications of non-uniform Fourier transform is presented.

Keywords: Multidimensional NMR, sparse sampling, non-linear sampling, biomolecular NMR

2

1 Introduction

Nuclear Magnetic Resonance (NMR) spectroscopy is one of the most important tools in structural studies of chemical compounds, ranging from small molecules up to medium-sized proteins. The NMR spectra provide valuable information about molecular structure, interactions and dynamics. However, there is still a need for more robust and more effective methods of acquisition and processing of NMR data.

The early Nuclear Magnetic Resonance (NMR) spectroscopy utilized the Continuous Wave (CW) detection technique. It was based on continuous sweeping of the B_0 field strength or the frequency of electromagnetic wave, through the resonance conditions of nuclei in the assumed spectral range. The main drawbacks of CW detection were low sensitivity and loss of time needed for sweeping through empty spectral regions. The breakthrough in NMR spectroscopy was the development of pulse excitation for generation of the Free Induction Decay (FID) signal, and the observation that the time dependent FID signal and the NMR spectrum can be converted one into another by applying the Fourier transform (FT) [1]. This method greatly shortened spectral acquisition times and enabled the development of thousands of pulse sequences for numerous emerging applications. New experimental methods allowed the more accurate determination of parameters which were earlier difficult to obtain or inaccessible.

Despite considerable progress in the field, NMR spectroscopy still has the two significant limitations: the intrinsically low sensitivity, due to the low Boltzmann polarization of nuclear spins in thermal equilibrium, and the low dispersion of observed frequencies, due to a small differences in a nuclear shielding by surrounding electrons for nuclei of the same kind. The first problem is continuously circumvented by the technological developments, i.e. construction of higher field magnets, cryogenically cooled probe-heads and pre-amplifiers, modern electronics, cleaner RF sources, and recently, introduction of Dynamic Nuclear Polarization (DNP) technique enabling sensitivity gain of even two orders of magnitude. The new generations of NMR spectrometers feature higher sensitivity and allow studies of large molecules at lower concentration. The

problem of resolving overlapped resonances is more severe. Even spectra of simple molecules often exhibit a peak overlap. Additionally, the assignment of signal frequencies to the respective nuclei could be difficult and sometimes impossible. To some extent, in simple cases, the problem can be solved by employing a higher magnetic field, but the general approach to resolve the resonances is to spread them in a different frequency dimensions of multidimensional spectra. The idea was practically implemented by the indirect sampling of the spins evolution and referred to as the two-dimensional NMR spectroscopy [2, 3]. This development not only allowed resolution of individual peaks by introducing additional spectral dimensions, but also facilitated spectral assignment by detecting groups of mutually interacting nuclei which give rise to correlation peaks. The application of multiple polarization transfer revealed other important aspects of multidimensional spectroscopy: sensitivity enhancement by excitation and observation of FID signal of sensitive, high-γ spins, and observation of directly undetectable multiple quantum coherences. At the beginning, the two-dimensional NMR techniques were demonstrated to be useful for examination of small organic molecules. Soon, homonuclear 2D NMR experiments were successfully applied for studies of biological macromolecules in solution [4]. Later, with increasing availability of isotopically enriched proteins, significant improvement was achieved by introduction of triple-resonance three- and four-dimensional experiments utilizing scalar couplings for polarization transfers [5-9]. However, due to the reasons given below, acquisition of multidimensional NMR spectra with a sufficient resolution in all frequency dimensions can be an extremely time consuming task. The indirect sampling of spin coherences evolution, the key concept of multidimensional NMR experiments, is realized in a parametric way. This means: to sample a point of indirect time space a specific delay (or delays) in a pulse sequence should be set to achieve desired evolution time, and then one directly observed FID signal is acquired. As a consequence, in order to acquire a multidimensional spectrum one needs to record many single FID signals. The overall measurement time grows rapidly with a number of indirectly sampled dimensions and a desired resolution. A conventional *N-*

dimensional experiment requires acquisition of $2^{N-1} \cdot k_1 \cdot k_2 \cdot ... \cdot k_{N-1}$ single FID signals (where $k_i = sw_i \cdot t_{\max i}$, is the number of points in the ith dimension, sw_i and $t_{max\,i}$ are the required spectral width and maximum evolution time respectively, and the 2^{N-1} is the number of components needed for quadrature detection). The conventional sampling is performed with points placed on a Cartesian grid. In each dimension, spacing between points is related to the expected range of frequencies by the Nyquist Theorem (see Section 3.2.4). Thus, fulfilling the Nyquist theorem implicitly limits the maximum evolution time and, therefore, the obtainable resolution for the given duration of experiment. In a case of directly detected dimension the above limitation is insignificant. Here, data points are successively sampled by conversion of a voltage in a receiver circuit into numbers reflecting signal amplitude. The acquisition of a whole signal has to be performed in one step, and it usually takes from a fraction of second to a few seconds, until signal decays below the noise level. This does not significantly extend the experiment duration. Thus, the best possible resolution is almost always achievable at no additional cost. Moreover, in modern NMR spectrometers usually oversampling is employed, i.e. more points than necessary are sampled in order to improve spectral dynamic range and enable digital filtering [10].

The sampling limitations have more severe consequences in the case of indirectly sampled dimensions, where acquisition of each sampling point takes up to a few seconds. Even in 3D NMR experiments of proteins, featuring relatively fast transverse relaxation, it is almost impossible to reach the natural (determined by relaxation) line width in a reasonable experimental time. Limited experiment duration causes signal truncation and results in broadened spectral peaks, according to the Fourier Uncertainty Principle [11].

The problem of sampling requirements in multidimensional NMR is becoming relatively more severe with increasing B_0 fields. The stronger magnetic field increases proportionally separation between resonances, however, in the same time it broadens of spectral regions of interest. Hence, an x-fold increase in B_0, causes the necessity of x^{N-1}-fold extension of time required for N-dimensional experiment, in order to preserve the peak width. This effect,

although usually of minor importance for 2D experiments, became significant for the larger number of dimensions.

In the recent decade many approaches were proposed to overcome the sampling limitation problem. The most straightforward one is the modification of pulse sequences to allow increased repetition rate of FID signal acquisition, which leads to reduction of the experiment time [12-14]. It was also shown that the spatial encoding of spectral frequencies can be employed for measurement of multidimensional spectra in a single scan [15-20]. However, most of the efforts to accelerate acquisition of multidimensional NMR spectra were dedicated to the reconstruction of so called "sparsely sampled" spectra, i.e. with less data points than required by Nyquist condition. Both experiment duration and desired resolution can be optimized by the use of sparse sampling. The simplest version of sparse sampling is a straightforward signal truncation. In such a case it is possible to attempt signal extrapolation using linear prediction [21] or filter diagonalization method [22-24]. Enhanced spectral resolution can be also achieved from relatively highly truncated data sets employing Covariance Spectroscopy [25-30], and some variants of the maximum entropy method [31]. Another simple approach to undersampling is increasing of the distance between points which leads to shortened experiment duration, at the expense of peak aliasing. Thus, if chemical shifts are known from other experiments, assignment of cross-peaks is still possible [32]. The sparse sampling can also be applied in order to extend the sampled space in several ways. Among them, two are of the particular importance: sampling at constant intervals, but along the radius in a time domain [33]; or randomly [34]. The former option is utilized in projection spectroscopy, and requires the algebraic decoding of peak frequencies [35-39], or the reconstruction of multidimensional spectrum [40-43]. The latter enables to reconstruct a fully-dimensional spectrum featuring improved resolution and/or acquired faster than conventionally. The sparsely and randomly sampled data sets can be processed using FT [44-46], maximum entropy [47-49] or multidimensional decomposition [50-52] methods.

In this review we will focus on applications of FT to processing of non-uniformly (sparsely) data sets devoted to the reconstruction of high-resolution multidimensional NMR spectra.

2 Fourier Transform – basics

2.1 Definition

Fourier Transform (FT) is a mathematical operation, that converts function *s(t)* into function *S(f)* according to the formula:

$$S(f) = \int_{-\infty}^{+\infty} s(t) \cdot e^{-2\pi i f t} \, dt \qquad (1)$$

Which, for convenience, may be denoted as a linear operator "*FT*" acting on *s(t)*:

$$S(f) = FT[s(t)] \qquad (2)$$

Function $e^{-2\pi i f t}$ that multiplies signal *s(t)* is often referred to as a *transform kernel*.
Both *t* and *f* are real variables, while *s(t)* and *S(f)* may be complex in general. In many fields of signal processing (including NMR spectroscopy), the two variables correspond to *time* and *frequency* domains. Function *s(t)* is a time-domain signal recorded in the experiment. Function *S(f)* is its *frequency representation*, i.e. it shows how a signal can be decomposed into oscillatory functions of frequencies *f*. Knowing frequency representation of a signal, one can retrieve *s(t)* by applying Inverse Fourier Transform (IFT):

$$s(t) = \int_{-\infty}^{+\infty} S(f) \cdot e^{2\pi i f t} \, df \qquad (3)$$

Hence, *s(t)* and *S(f)* are equivalent representations of a signal and are often referred to as a *Fourier pair*. For the simplest infinite oscillatory signal of frequency ν the Fourier pair is:

$$e^{2\pi i \nu t} \quad \xrightarrow{FT} \quad \delta(f - \nu)$$
$$\xleftarrow{IFT}$$

$$(4)$$

where $\delta(f - \nu)$ is the Dirac delta, and can be informally thought of as an infinitely narrow and infinitely high peak centered at ν:

$$\delta(f - \nu) = \begin{cases} +\infty & \text{for } f = \nu \\ 0 & \text{for } f \neq \nu \end{cases}$$

$$(5)$$

The result can be explained by the orthogonality of oscillatory exponentials, that are *basis functions* for FT i.e.:

$$\int_{-\infty}^{+\infty} e^{2\pi i \nu t} \cdot e^{-2\pi i f t} \, dt = \begin{cases} +\infty & \text{for } f = \nu \\ 0 & \text{for } f \neq \nu \end{cases}$$

$$(6)$$

Thus, as the result of FT one obtains function that reaches high values for coordinates corresponding to frequencies present in a signal. This function, called *spectrum* is of particular interest, especially in scientific tasks. Representing oscillatory time-domain signal as a peak in frequency domain often provides better insight into physical phenomena, as discussed in the next section. Description of measured signals based on complex numbers may be quite confusing and requires brief explanation. Notably, complex signal is artificially constructed from actually measured real-valued signals of the same frequencies and amplitudes, but shifted in phase by $\frac{\pi}{2}$:

$$s(t) = s_{\cos}(t) + i s_{\sin}(t)$$

$$(7)$$

e.g. $s(t) = e^{2\pi i \nu t}$ consists of $s_{\cos}(t) = \cos(2\pi \nu t)$, and $s_{\sin}(t) = \sin(2\pi \nu t)$.
Equivalently, one may use two-dimensional vector to describe complex signal:

$$s(t) = \begin{pmatrix} s_{\cos}(t) \\ s_{\sin}(t) \end{pmatrix}$$

(8)

This notation will be used in discussion of multidimensional FT in the next section. Variants of FT featuring real-valued kernel i.e. *Cosine FT* (Cos-FT) and *Sine FT* (Sin-FT), can be defined as:

$$S_{\cos}(f) = \int_{-\infty}^{+\infty} s(t) \cdot \cos(2\pi f t)\, dt$$

(9a)

$$S_{\sin}(f) = \int_{-\infty}^{+\infty} s(t) \cdot \sin(2\pi f t)\, dt$$

(9b)

Again, for convenience, operator notation can be introduced:
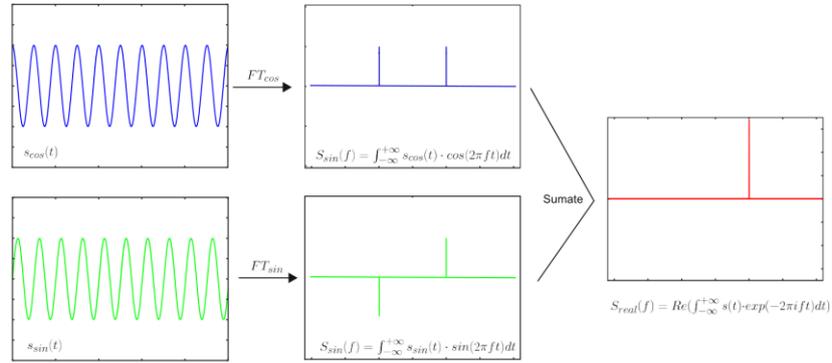
$$S_{\cos}(f) = FT_{\cos}[s(t)]$$

(10a)

$$S_{\sin}(f) = FT_{\sin}[s(t)]$$

(10b)

Complex FT of a complex signal can be described as a sum of *Cosine FT* and *Sine FT*:

$$S(f) = FT_{\cos}[s_{\cos}(t)] + FT_{\sin}[s_{\sin}(t)] - i(FT_{\sin}[s_{\cos}(t)] - FT_{\cos}[s_{\sin}(t)])$$ (11)

This notation allows easy visualization of the essence of complex FT (see Figure 1).

Fig. 1. Idea of complex FT. Two signals of the same frequency and amplitude, shifted in phase by π/2 $s_{cos}(t)$ and $s_{sin}(t)$ are transformed with cosine and sine FT and added. This may be described as one complex operation on one complex signal $s(t)=s_{cos}(t)+is_{sin}(t)$.

2.2 Multidimensional FT

Fourier Transform can be extended to N dimensions:

$$S(\vec{f}) = \int_{\Re^N} s(\vec{t}) \cdot \left[\begin{pmatrix} \cos(2\pi f_1 t_1) \\ \sin(2\pi f_1 t_1) \end{pmatrix} \otimes \begin{pmatrix} \cos(2\pi f_2 t_2) \\ \sin(2\pi f_2 t_2) \end{pmatrix} \otimes ... \otimes \begin{pmatrix} \cos(2\pi f_N t_N) \\ \sin(2\pi f_N t_N) \end{pmatrix} \right] d\vec{t}$$

(12)

Where $\vec{f}, \vec{t}$, are N-dimensional vectors:

$$\vec{f} = (f_1, f_2, ..., f_N)$$
$$\vec{t} = (t_1, t_2, ..., t_N)$$

It is noteworthy, that transform kernel is represented by direct product ($\otimes$) of one-dimensional complex functions. The kernel is thus an 2^N-dimensional vector:
One can represent both signal $\hat{s}(\vec{t})$ and spectrum $\hat{S}(\vec{f})$ in a similar fashion:

$$\hat{s}(\vec{t}) = s(t_1) \otimes s(t_2) \otimes ... \otimes s(t_N)$$

$$\hat{S}(\vec{f}) = S(f_1) \otimes S(f_2) \otimes ... \otimes S(f_N)$$

where, $S(f_i) = FT[s(t_i)]$

First element of $\hat{S}(\vec{f})$ corresponds to the real part of a spectrum.

The FT of simplest multidimensional signal is thus multidimensional delta function, i.e.:

$$\begin{pmatrix} \cos(2\pi v_1 t_1) \\ \sin(2\pi v_1 t_1) \end{pmatrix} \otimes \begin{pmatrix} \cos(2\pi v_2 t_2) \\ \sin(2\pi v_2 t_2) \end{pmatrix} \otimes ... \otimes \begin{pmatrix} \cos(2\pi v_N t_N) \\ \sin(2\pi v_N t_N) \end{pmatrix} \xrightleftharpoons[IFT]{FT} \delta(f_1 - v_1, f_2 - v_2, ..., f_N - v_N)$$

(13)

where

$$\delta(f_1 - v_1, f_2 - v_2, ..., f_N - v_N) = \delta(f_1 - v_1) \otimes \delta(f_2 - v_2) ... \otimes \delta(f_N - v_N).$$

Again, signal frequency (frequencies) is clearly visualized in the spectral domain as a "peak" centered at $(v_1, v_2, ..., v_N)$. Its position informs about correlated frequencies present in the multidimensional signal, which is usually the most essential experimental information.

2.3 FT – two basic features

In the end of this section, we would like to mention two of the most important features of FT, which will be helpful in the analysis of specific case of NMR signal. These are:

1. Linearity:

$$FT[\alpha f(t) + \beta g(t)] = \alpha FT[f(t)] + \beta FT[g(t)]$$

(14)

2. Convolution Theorem:

$$FT[f(t) \cdot g(t)] = FT[f(t)] * FT[g(t)]$$

(15)

where $*$ denotes convolution, defined as (see also Figure 2a,b):

$$u(x) * v(x) = \int_{-\infty}^{\infty} u(x) \cdot v(y - x) dy$$

(16)

Two above features of FT will help us to evaluate how simple manipulations of the signal, like multiplication and addition affect its spectrum. Notably, using only these two kinds of operations allows to change from monochromatic, non-decaying, perfectly continuous and infinite signal discussed above to actually measured NMR signal.

3 Fourier Transform of the NMR signal

3.1 Perfect FID

Free Induction Decay (FID) signal is a time-domain function resulting from NMR measurement. Although it is quite complicated, one can easily separate its features and discuss how they manifest themselves in frequency representation, i.e. a spectrum. These features are: relaxation, signal amplitude and multiple components (see Figure 2).

3.1.1 Relaxation

NMR signal decays exponentially with time (or with "times", in multidimensional case). This can be represented by element-wise multiplication of a signal $\hat{s}_0(\vec{t}) = \left(e^{-i2\pi\nu_1 t_1}, e^{-i2\pi\nu_2 t_2}, ... e^{-i2\pi\nu_N t_N} \right)^T$ by decaying exponential:

$$\hat{s}_1(\vec{t}) = \hat{s}_0(\vec{t}) \circ \left(e^{-R_1 t_1}, e^{-R_2 t_2}, ... e^{-R_N t_N} \right)^T$$

(17)

The FT of $\hat{s}_1(\vec{t})$ is, according to statements from the section 2.3, a

convolution of multidimensional peak $\delta(f_1 - v_1, f_2 - v_2, ..., f_N - v_N)$ (i.e. $FT[\hat{s}_0(\vec{t})]$) and Lorentzian function (being FT of a decay function):

$$FT[\hat{s}_1(\vec{t})] = \delta(f_1 - v_1, f_2 - v_2, ..., f_N - v_N) * \left(\frac{1}{R_1 + if_1}, \frac{1}{R_2 + if_2}, ... \frac{1}{R_N + if_N} \right)$$

(18)

The result is Lorentzian peak centered at $(v_1, v_2, ..., v_N)$.

3.1.2 Amplitude

NMR signal has certain amplitude. This can be represented by multiplying $\hat{s}_1(\vec{t})$ by some constant A:

$$\hat{s}_2(\vec{t}) = A \cdot \hat{s}_1(\vec{t})$$

(19)

Obviously, multiplying signal by constant is equivalent to multiplication of spectrum by the same constant:

$$FT[A \cdot \hat{s}_2(\vec{t})] = A \cdot FT[\hat{s}_1(\vec{t})]$$

(20)

3.1.3 Multiple components

NMR signal consists of multiple components, corresponding to groups of equivalent spin systems. Each of the components has its own amplitude and relaxation parameters:

$$\hat{s}_3(\vec{t}) = \sum_i \hat{s}_2^i(\vec{t})$$

(21)
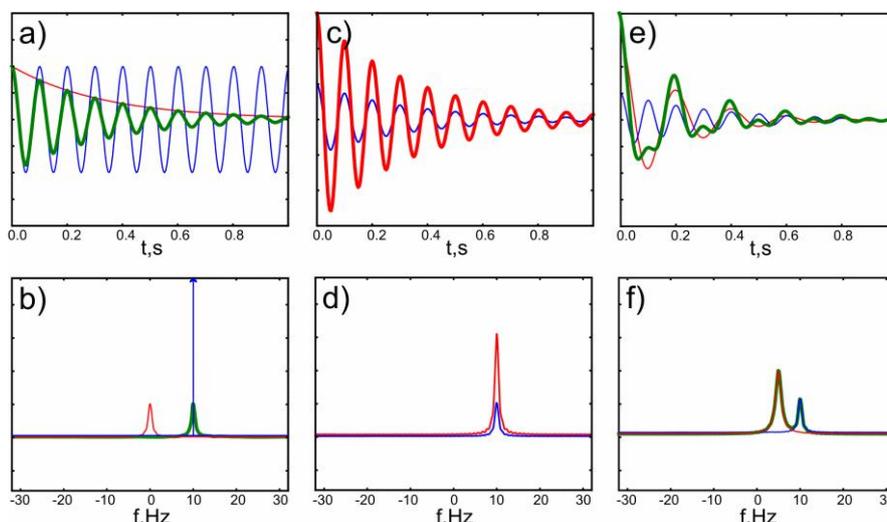
Fig. 2. The main features of FID signal and its spectrum: (a) Relaxing NMR signal (bold line) is product of decaying exponential function and oscillatory function (thin lines). (b) Spectrum of relaxing NMR signal (bold line) is convolution of Lorentzian function and a delta peak (thin lines). (c) Relaxing NMR signal of some amplitude A (bold line) is decaying sinusoid (thin line) multiplied by constant A. (d) Spectrum of relaxing NMR signal of some amplitude A (bold line) is Lorentzian peak (thin line) multiplied by constant A. (e) Multi-component NMR signal (bold line) is a sum of decaying components of different amplitudes (thin lines). (f) Spectrum of multi-component NMR signal (bold line) is a sum of spectra of individual components (thin lines).

To summarize, as a model of multidimensional FID signal one can use an oscillatory function consisting of multiple, decaying components of various amplitudes. Spectrum of such signal is built of Lorentzian peaks centered at frequency coordinates corresponding to component frequencies. Peak heights are proportional to component amplitudes in time domain. Peak half-widths are inverse of signal decay rates.
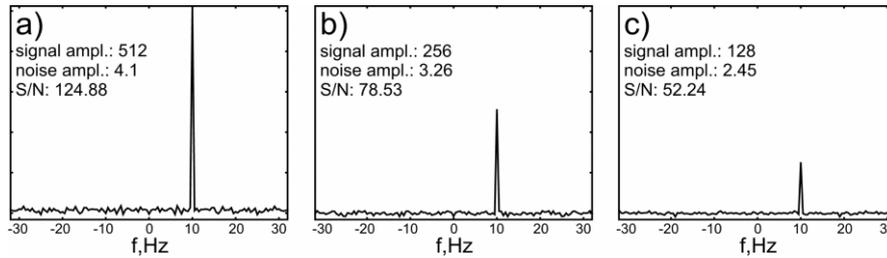
3.2 Measured FID

Above model gives an idea, how *perfect* signal and its spectrum look like.
The *real* output of an NMR experiment is quite far from the model. Three factors are most important here, namely: noise, finite measurement time and sampling.

3.2.1 Noise

NMR signal contains some random noise $\varepsilon(t)$. Its spectrum is thus the sum of two FTs:

$$\hat{S}(\vec{f}) = FT[\hat{s}_3(\vec{t})] + FT[\varepsilon(\vec{t})]$$

(22)

Fig. 3. Peak amplitude, noise level and signal-to-noise ratio for spectrum of non-decaying signal of frequency 10 Hz, sampled with: a) 512, b) 256, c) 128 pts. Noise is white and Gaussian.

Assuming that the noise is *white* and *Gaussian* i.e. its amplitude is independent of frequency and described by Gaussian distribution, the signal to noise ratio is proportional to the square root of the number of measurements (see Figure 3).

3.2.2 Finite measurement time

Obviously, maximum time of spin evolution cannot be infinite (which would be pointless anyway, because of relaxation). This limit can be represented by multiplying signal by step function $\Pi(t_1, t_2 ... t_m)$:

$$\Pi(t_1, t_2 ... t_m) = \begin{cases} 1 & t_i < t_{i\max}, i = 1, ... m \\ 0 & otherwise \end{cases}$$

(23)

where t_{imax} is a maximum evolution time set in i[th] spectral dimension. Signal $\hat{s}_{t\max}(\vec{t})$, i.e. time-limited noiseless NMR signal can be described as:

$$\hat{s}_{t\max}(\vec{t}) = \hat{s}_3(\vec{t}) \cdot \Pi(t_1, t_2 \ldots t_m) \tag{24}$$

The effect of multiplication in time domain is, according to Convolution Theorem, the convolution in frequency domain:

$$FT[\hat{s}_{t\max}(\vec{t})] = FT[\hat{s}_3(\vec{t})] * FT[\Pi(t_1, t_2 \ldots t_m)] \tag{25}$$

The FT of step function is a *sinc* function of width inversely proportional to t_{imax} (see Figure 4):

$$FT\{\Pi(\vec{t})\} = \frac{\sin(2\pi f_1 \cdot t_{1\max})}{2\pi f_1 \cdot t_{1\max}} \frac{\sin(2\pi f_2 \cdot t_{2\max})}{2\pi f_2 \cdot t_{2\max}} \ldots \frac{\sin(2\pi f_N \cdot t_{N\max})}{2\pi f_N \cdot t_{N\max}} \tag{26}$$

Thus, finite acquisition time causes a convolution of NMR spectrum with *sinc* function. This manifests itself in peak broadening and presence of *sinc* "wiggles". The broadness of the NMR peak is thus dependent not only on relaxation rate but also on the maximum evolution time. Both effects correspond to Fourier Uncertainty Principle [53] stating that, in general, the "broadness" of time representation and frequency representation are inversely proportional to each other.
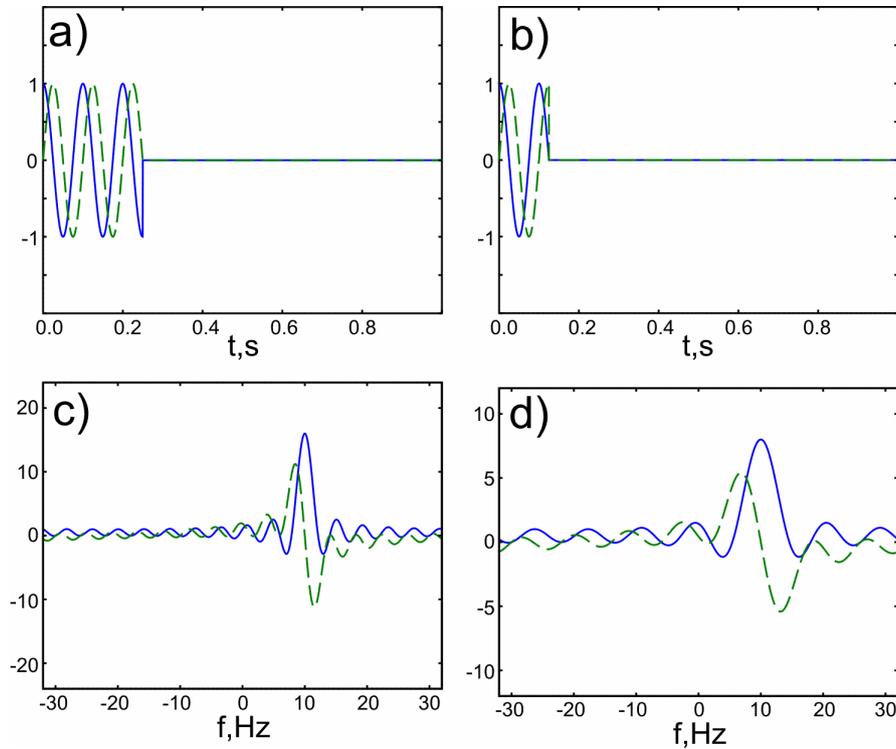
Fig. 4. Signal truncation and spectral line width (real and imaginary parts marked with solid and dashed lines appropriately). a) signal truncated to 250 ms b) signal truncated to 125 ms c) spectrum of a signal truncated to 250 ms – *sinc* function d) spectrum of a signal truncated to 125 ms – *sinc* function.

3.2.4 Sampling

NMR signal is measured in discrete manner i.e. sampled. This may be represented by *sampling function*, being a multidimensional train of *K* delta pulses:

$$III(t_1, t_2 ... t_m) = \sum_{k=1}^{K} \delta\left(t_1 - t_1^k, t_2 - t_2^k, ..., t_N - t_N^k\right)$$

(27)

Discrete sampling can be thus represented by multiplication of continuous signal by $III(t_1, t_2 ... t_m)$.

In general, one can distinguish between two kinds of sampling:

uniform (or conventional) i.e. with sampling coordinates $\left(t_1^k, t_2^k, ..., t_N^k\right)$ corresponding to full Cartesian grid and *non-uniform* i.e. with coordinates chosen arbitrarily, according to one of the sampling schedules (see section 4).

Conventional discrete sampling influences spectrum in quite straightforward way, because of simplicity of corresponding Fourier pair:

$$III\left(\frac{t_1}{\tau_1}, \frac{t_2}{\tau_2}...\frac{t_m}{\tau_m}\right)\xrightarrow[\text{IFT}]{\text{FT}} III\left(\tau_1 f_1, \tau_2 f_2 ... \tau_m f_m\right)$$

(28)

As a result, an infinite number of "copies" are produced in spectral domain, with the distance $\left(\frac{1}{\tau_1}, \frac{1}{\tau_2}...\frac{1}{\tau_m}\right)$ between "copies" (see Figure 5 and 6a,b). If the distance is greater than half of signal bandwidth, then copies do not overlap. This leads to well known Shannon-Nyquist Sampling Theorem [54], saying that spectrum of band-limited signal can be perfectly recovered from discrete samples if sampling frequency is at least twice higher than highest frequency present in the signal. Obviously, if spectrum is perfectly recovered, then *continuous* signal is recovered as well, meaning that discrete points are interpolated with *sinc* functions.

Usually, the sampling frequency is set basing on predicted spectral width of a signal. Assuming, that the frequency band is limited, one can take the central, low-frequency spectrum "copy" that is equivalent to perfect spectrum (see Figure 5b). Actually, this is done by default by standard FT algorithms, e.g. Fast Fourier Transform (*FFT*). In the case, when the prediction is wrong, i.e. sampling interval is higher than required by Shannon-Nyquist criterion, the "copies" of spectrum overlap, which leads to phenomenon known as *peak folding* or *aliasing*. This phenomenon manifests itself by presence of peaks at false frequency coordinates in the spectral region of interest.
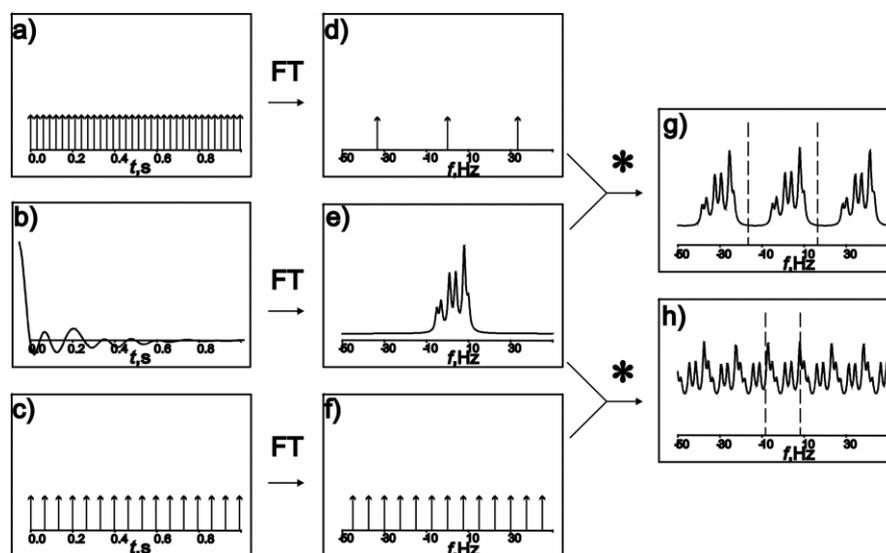
Fig. 5. Aliasing phenomenon. (a) Train of delta pulses representing sampling with τ=0.03 s (sampling rate 33.33 Hz); (b) Continuous signal - multi-component oscillatory function; (c) Train of delta pulses representing sampling with τ=0.066 s (sampling rate 15 Hz); (d) Train of delta pulses in frequency domain being FT of sampling schedule (a); (e) Spectrum of continuous signal (b); (f) Train of delta pulses in frequency domain being FT of sampling schedule (c); (g) Convolution of (d) and (e) corresponding to FT of signal (b) sampled according to (a). Properly sampled spectral bandwidth with central spectrum "copy" marked with dashed line. No aliasing; (g) Convolution of (f) and (e) corresponding to FT of signal (b) sampled according to (c). Properly sampled spectral bandwidth with central spectrum "copy" marked with dashed line. Aliasing due to insufficient sampling rate.

It should be noted, that upper limit for the sampling interval results in lower limit for peak width. This becomes especially significant in the multidimensional NMR experiments, where each sampling point takes few seconds of experimental time (see Scheme 1). Moreover, the requirements of regular sampling grow exponentially with the number of dimensions and despite hours- or days-long measurements natural, relaxation-determined peak widths are rarely obtained even in 3D spectra.

Coupling between peak width and number of sampling points (i.e. experimental time) is the main reason for the use of non-uniform sampling in NMR.

Expected spectral width

$\downarrow$ *Sampling Theorem*

Distance between points

$\downarrow$ *Given time of spectrometer use*
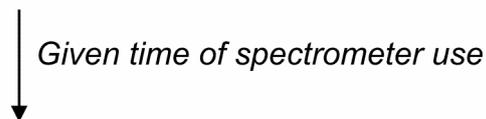
Limited maximum evolution time

$\downarrow$ *Convolution with „sinc"*
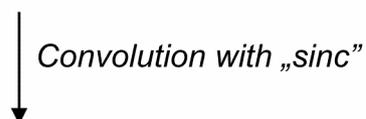
Line broadening (limited resolution)

Scheme 1. Scheme illustrating limitations associated with conventional sampling i.e. coupling between experimental time and line width.

For non-uniform sampling the Fourier Transform of sampling schedule (sometimes referred to as *Point Spread Function, PSF*) is not a simple train of delta pulses, but becomes more complex function (see section 4 and Figure 6). This, in general, leads to three conclusions:

1. For arbitrary, non-uniform sampling it is no longer possible to obtain spectrum that is equal to spectrum of continuous signal, even if it is strictly band-limited. Spectral artifacts, depending on the sampling schedule appear as a part of Point Spread Function.

2. Aliasing, however, is caused only by presence of sampling grid, from which sampling points are taken. For purely off-grid sampling aliasing does not appear (notably, NMR hardware allows very fine approximation of off-grid sampling). This means, that one can use non-uniform sampling to remove coupling between sampling rate and line

width and obtain high spectral resolution in relatively short experimental time (see Section 8).

3. The negative effect of non-regular sampling, i.e. presence of artifacts, is separate from all other spectral effects associated with various signal features e.g. relaxation, amplitude etc. and can be discussed independently. In other words, FT of irregularly sampled signal has the same features as FT of conventionally sampled signal, differing only in PSF.

4 Non-uniform sampling schemes

Conventional (Cartesian grid) sampling scheme is an obvious method of choice, when experimental time and/or line width expense is acceptable. However, when especially narrow spectral lines or high dimensionality are required, an irregular sampling should be employed. As stated above, it can make peak widths independent of experimental time. Nevertheless one should always remember about cost of irregularity i.e. introduction of spectral artifacts, whose pattern and level is dependent on a sampling scheme.
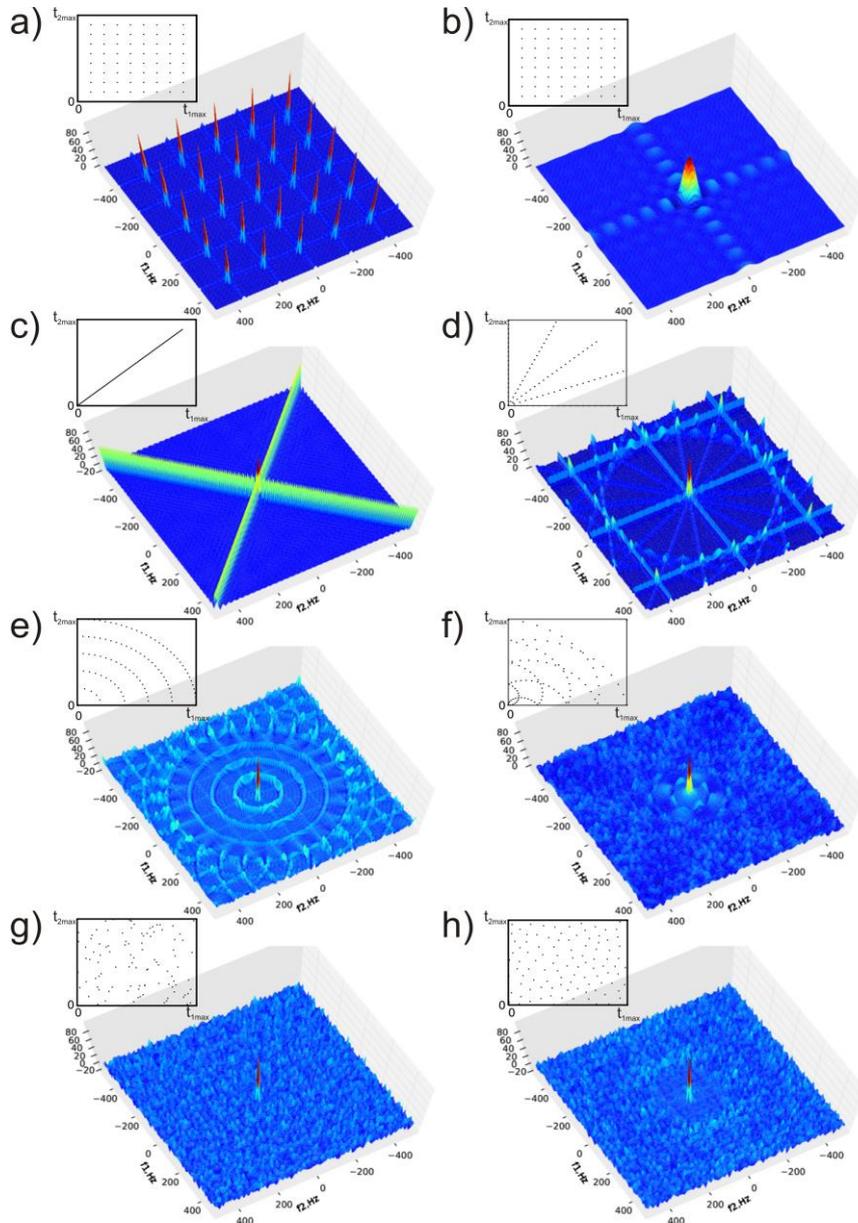
Fig. 6. Point Spread Functions of various sampling schemes (presented in the upper left part of each panel): a) conventional sampling ($t_{1max}=t_{2max}=50$ ms), b) conventional sampling ($t_{1max}=t_{2max}=5$ ms), c) radial sampling with one sampling line, d) radial sampling with five sampling lines, e) concentric rings sampling, f) spiral sampling, g) purely random sampling, h) Poisson disk sampling. Number of points in each sampling scheme is equal to 100. For panels c-h $t_{1max}=t_{2max}=50$ ms.

4.1 Radial sampling

Radial sampling was the first sparse sampling scheme introduced to NMR. Apart from FT [44, 55, 56], other techniques of data processing were proposed. These include: reduced dimensionality [57], projection-reconstruction [41], multi-way decomposition [58]. Radial sampling scheme consists of points placed on a set of lines in time domain (see Figure 6c,d). The coordinates of ith point, lying on j^{th} line can be described in polar coordinate system as:

$$t_1^{i} = i \cdot \Delta r \cdot \cos \psi^{j}$$

(29a)

$$t_2^{i} = i \cdot \Delta r \cdot \sin \psi^{j}$$

(29b)

PSF of radial distribution is a set of ridges (see Figure 6c,d). Each pair of ridges is FT of one sampling line and they are oriented at ψ^{j} and $\psi^{j} + \pi/2$ angles.

4.2 Concentric rings sampling

Concentric rings sampling was proposed by Coggins and Zhou [59]. Sampling scheme is depicted in Figure 6e. Importantly, the number of points situated on each ring increases linearly with a ring's radius (linearly increasing concentric ring sampling, LCRS). The coordinates of ith time point, lying on j^{th} ring are:

$$t_1^{i} = r \cdot \cos(i \cdot \Delta \psi^{j} + \psi_0^{j})$$

(30a)

$$t_2^{i} = r \cdot \sin(i \cdot \Delta \psi^{j} + \psi_0^{j})$$

(30b)

In LCRS ψ_0^{j} is the same for all rings. If this phase is chosen randomly for each ring independently, the scheme is called RLCRS (randomized LCRS). PSF of LCRS takes form of a set of ring-

shaped ridges, and for RLCRS this pattern is slightly disturbed, covering the spectral space more evenly.

4.3 Spiral sampling

Point coordinates in spiral sampling scheme [44] are defined as:

$$t_1^i = i \cdot \Delta r \cdot \cos(i \cdot \Delta \psi^j)$$

(31a)

$$t_2^i = i \cdot \Delta r \cdot \sin(i \cdot \Delta \psi^j)$$

(31b)

Notably, both radial and LCRS schemes can be considered as special cases of spiral sampling scheme. PSF of this distribution is the combination of two above PSFs. The artifacts form ring-shaped ridges, but the intensity is not constant along the rings, but varies with angle (see Figure 6f) [60].

4.4 Random sampling

As can be clearly seen from the above examples, regularity in time domain results in regularity in frequency domain. This suggests that very irregular, random sampling schemes can be particularly useful (see Figure 6g). In this case the artifacts are "spread" evenly over the spectral space, their level is thus reduced comparing to more regular sampling schemes and consequently it is less probable for false peaks to come up.

Optimization of such purely random distribution can be done by introducing certain constraints which protect against choosing one point too close to another. A few algorithms of generation of such semi-random sampling schemes were investigated [60, 61]. Among them Poisson disk sampling was found to be the most optimal (see Figure 6h). It directly assumes a minimal distance between time points. The artifact level is not as even as in purely random case. It is lower in the vicinity of the peaks. Moreover, by slight

modification of the restraints one can adjust the shape of the "clean" region to spectral widths or to compensate for different maximum evolution times in different dimensions. Another variant of distance-restrained sampling, referred to as Poisson-gap sampling, was presented by Hyberts and coworkers and used with forward maximum entropy processing [62].

4.5 Weighted samples and weighted probability

In a conventional case, usually certain weighting function is applied in each dimension for improving signal-to-noise ratio or reduce effect of signal truncation (*sinc* "wiggles"). This procedure is called Weighted Samples (WS) and can be performed also for random sampling schemes. In the case of irregular sampling, however, also an alternative solution can be employed. Instead of applying weighting function to a sampled signal, the function may be used as a probability distribution during generation of randomized sampling scheme. Such approach is referred to as Weighted Probability (WP). The two procedures (WS and WP) result with spectra of the same line shape and S/N for non-decaying signal (i.e. if S/N is constant in time) [63], see Figure 7. However in the case of a real, relaxing and noisy FID signal more effective is the WP method, as more points of higher S/N (from the beginning of FID) are measured [60].
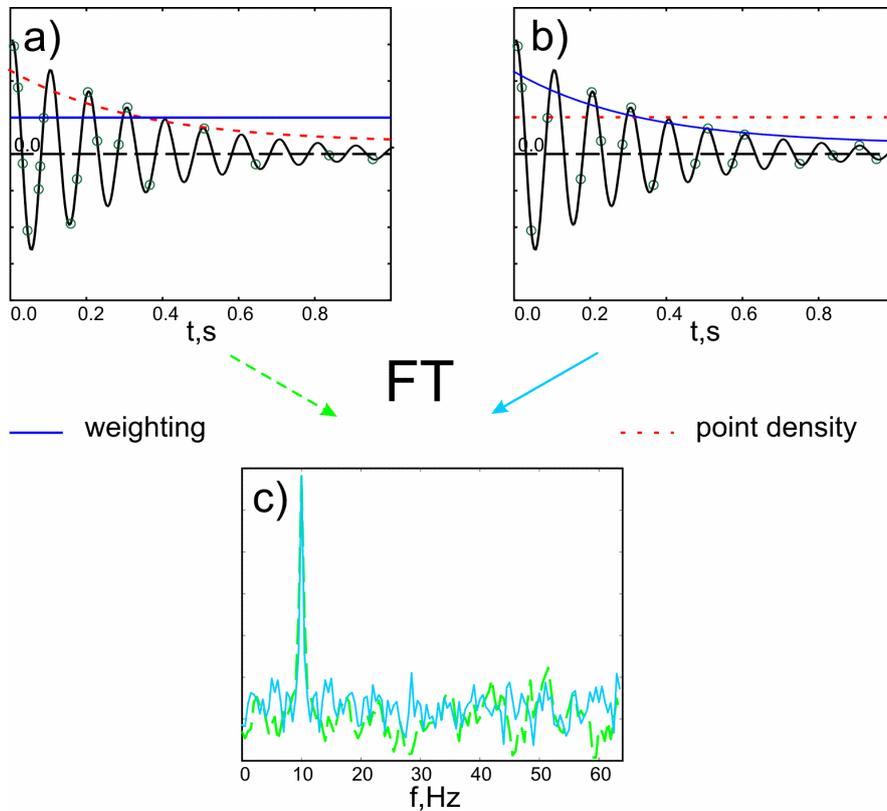
Fig. 7. Comparison of simulated spectra obtained using Weighted Probability (exponential distribution, solid line) and Weighted Samples (exponential weighting, dashed line) methods. The signal was simulated without a thermal noise. Both methods give spectra with the same signal to artifact ratio and line widths.

5 Methods of integration

FT is an integral operation. In a real case the transformed function is discrete, thus the integral has to be replaced by a sum. Therefore, for irregular points distributions unequal distances between sampling points can be taken into account by applying certain weights. In analogy to 1D numerical integration employing a rectangular or trapezoidal rule, for multidimensional case one can obtain weights by Voronoi tessellation [60, 64] or Delaunay triangularization [46]. These methods, although helpful in the case of polynomials

integration or rational functions, are rather unsuitable for rapidly oscillating functions. When the sampling density is lower than density determined by Sampling Theorem (i.e. always, when non-uniform sampling is justified) introduction of mentioned weights diminishes the signal-to-artifact ratio (see Figure 8). Therefore, for oscillatory functions more appropriate is simple summation i.e. Monte Carlo integration [65, 66]. It does not affect signal-to-noise ratio and the result converges to the exact value with $\sqrt{n}$.
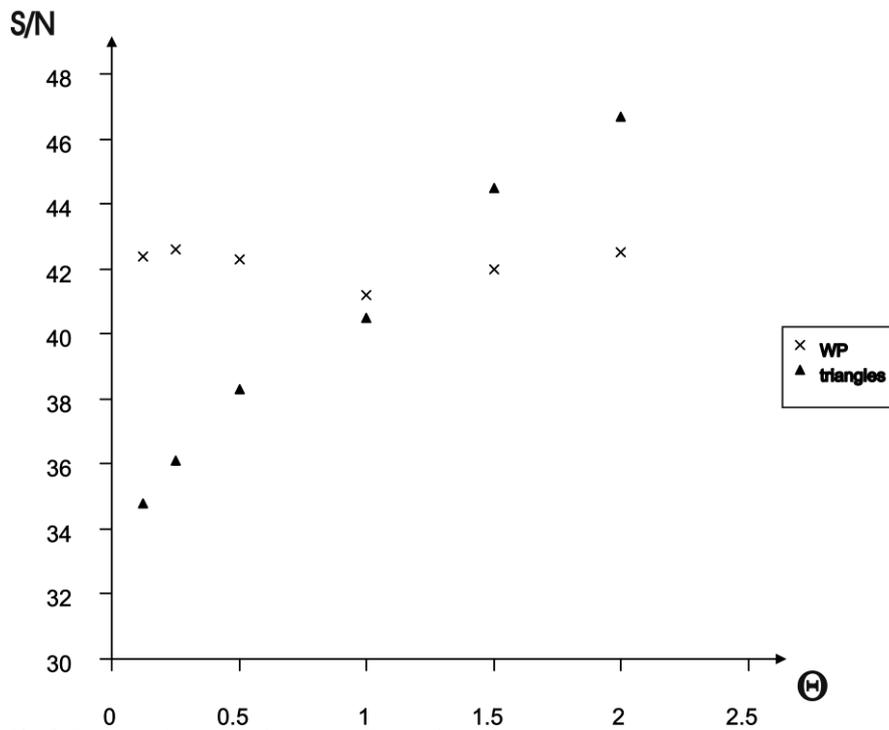
Fig. 8. The plot of spectral signal to artifact ratio of simulated spectrum $f(t_1, t_2) = \exp(-2\pi i_1 \cdot 300Hz \cdot t_1 - 2\pi i_2 \cdot 300Hz \cdot t_2 - 50Hz \cdot t_2)$ in function of relative density of time domain points $(\Theta = \rho / \rho_N)$ comparing: WP method and surface integration procedure (512 evolution time points of Gaussian PDF: $\exp(-t^2 / \sigma^2), \sigma = 0.5$). Spectral widths and maximum evolution times were equal $sw_1 = sw_2$, $t_{1max} = t_{2max} = t_{max} = 0.02$ s. Θ was changed by varying both spectral widths (and ρ_N consequently) keeping constant number of points and evolution time surface t_{max}^2 (and ρ consequently). Reprinted with permission from Ref. [46].

6 Sparse sampling and FT as a linear algebra issue

One-dimensional FT of a sampled signal (or its spectrum), may be thought of as a solution to a system of equations:

$$\frac{1}{\sqrt{N}}\begin{bmatrix} e^{i2\pi f^1 t^1} & e^{i2\pi f^2 t^1} & \dots & e^{i2\pi f^m t^1} \\ e^{i2\pi f^1 t^2} & e^{i2\pi f^2 t^2} & \dots & e^{i2\pi f^m t^2} \\ \dots & \dots & \dots & \dots \\ e^{i2\pi f^1 t^n} & e^{i2\pi f^2 t^n} & \dots & e^{i2\pi f^m t^n} \end{bmatrix} \begin{bmatrix} S(f^1) \\ S(f^2) \\ \dots \\ S(f^m) \end{bmatrix} = \begin{bmatrix} s(t^1) \\ s(t^2) \\ \dots \\ s(t^n) \end{bmatrix} \tag{32}$$

Or, more briefly:

$$\hat{A}\vec{S} = \vec{s} \tag{33}$$

Where $\hat{A}$ is an inverse FT matrix with number of rows n equal to the number of time points and number of columns m equal to the number of frequency points, $\vec{S}$ is an m-element vector representing spectrum and $\vec{s}$ represents vector of n signal samples:

$$A_{ij} = \frac{e^{i2\pi f^j t^i}}{\sqrt{N}} \tag{34a}$$

$$S_i = S(f^i) \tag{34b}$$

$$s_j = s(t^j) \tag{34c}$$

Thus, the usual spectral processing task is to find unknown $\vec{S}$ that agrees with known $\vec{s}$ (fulfills the system of equations). The possible situations are:

1. $n=m$ and matrix $\hat{A}$ is full rank. Then, the system of equations has unique solution and it may be obtained by multiplying both sides by FT matrix, which is Hermitian transpose (conjugate and

transpose) of matrix $\hat{A}$:

$$\vec{S} = \left(\hat{A}^*\right)^T \vec{s}$$

(35)

In the case of equally spaced sampling, matrix $\hat{A}$ is highly symmetric and FFT algorithms may be employed to reduce computational time.

The number of samples taken from the signal (*n*) determines the number of frequency points that are possible to be determined. Adding zero-valued, "artificial" sampling points at the end of the signal allows to calculate increased number of spectral points. This procedure (known as *zero filling*) is equivalent of interpolation in the spectral domain [67].

2. *n>m.* The system of equations is overdetermined and strictly speaking there is no solution. However, the number of equations may be reduced and solution can be obtained with additional gain on signal-to noise ratio. This is achieved by various digital filtering techniques. The situation corresponds to *oversampling* and in practice exists only in directly detected signal.

3. *n<m.* The system of equations is underdetermined and there are many possible solutions. This corresponds to sparse, non-uniform sampling.

Among the spectral vectors, that fulfill the system of equations, there is an optimal one i.e. spectrum of a signal sampled in an uniform manner. Finding it, however, is not a simple task (even if thermal noise could be neglected). Many approaches were presented, differing in type of constraints, that limits the number of solutions. Some of these include:

Maximum Entropy Methods - the solution with highest entropy is found. Various "entropy" functions were used in the past [49, 68].

Integration of frequency and time domain information [69] by assuming that some of the frequency points are equal to zero.

l_1-norm minimization, the solution with smallest l_1-norm (sum of

absolute values of spectral points) is found [70, 71]. It was proved recently that for signals featuring "dark" spectra (small number of non-zero frequencies) l_1-norm should lead to optimal solution by convex minimization [72]. This approach, has been successfully employed in many branches, including MRI [73, 74]. Reconstruction of NMR spectra with this method is very costly (days-long calculations [70]) or suffers from spectral disturbances [71].

Interpolation or gridding of sparse dataset may help to recover missing data points and use conventional FT processing. This, however, may lead to significant disturbances if simplest, polynomial interpolation is [75]. More advanced gridding techniques are helpful here [76].

Non-uniform FT (nuFT) employing equation identical to (35), but with non-quadratic matrix $\hat{A}$ (or with full $m \times m$ matrix $\hat{A}$ and zeros at non-sampled points of m-long vector $\vec{s}$). The obtained solution features minimum l_2-*norm* (power), which can be easily proved, considering that FT is an unitary operation and thus l_2-norms of signal and spectrum are equal (Parseval's theorem). Although, the solution is not the optimal one, the processing is fast and was successfully employed in many applications [77-82]. Moreover, the spectrum can be additionally improved by application of various artifact-cleaning algorithms [60, 81, 83].

7 Suppression of sampling artifacts in FT spectra

7.1 The principle of CLEAN algorithm

7.1.1 The model of "dark" spectrum

As mentioned in Section 6, in the case of sparse sampling there is insufficient data to uniquely determine the Fourier representation of the measured signal. Therefore, sampling artifacts observed in nuFT

spectra can be regarded as an unavoidable consequence of missing data. However, more accurate spectral estimates can be obtained by incorporating *a priori* knowledge about the nature of sampled signal. For a certain class of signals it might be assumed that the continuous Fourier spectrum consists of small number of well-localized components (peaks) and relatively weak flat (frequency independent) noise:

$$\hat{S}(\vec{f}) = \sum_i \hat{S}_i(\vec{f}) + \varepsilon(\vec{f})$$

(36)

This general model, usually referred to as "dark" spectrum, allows a variety of reconstruction methods to be employed (see Section 6). The CLEAN algorithm, proposed originally for the reconstruction of two-dimensional maps in radio-astronomy [84], utilizes essentially the same signal properties. It is noteworthy that the "dark spectrum" model is especially well-suited to multidimensional NMR spectroscopy.

7.1.2 Description of the CLEAN procedure

The starting point of the procedure is a discrete FT spectrum

$$\hat{S}(\vec{f})^{(0)} = \hat{S}(\vec{f}) = FT[s(\vec{t}) \cdot III(\vec{t})]$$

(37)

which is a convolution of continuous spectrum with the FT of sampling function (Point Spread Function). The latter is usually termed "dirty mask" in this context.
The aim of the procedure is to identify well-localized sources of artifacts present in FT spectra. Intuitively, one may suppose, that they can be found by computing the convolution of FT spectrum and the PSF:

$$\hat{C}(\vec{f}) = \hat{S}(\vec{f}) * III(\vec{f})$$

(38)

Following the convolution theorem (see Section 2.3) one obtains that this convolution is the FT spectrum itself:

$$\hat{C}(\vec{f}) = \hat{S}(\vec{f}) * III(\vec{f}) = FT[\hat{s}(\vec{t}) \cdot III(\vec{t}) \cdot III(\vec{t})] = FT[\hat{s}(\vec{t}) \cdot III(\vec{t})] = \hat{S}(\vec{f})$$
$$(39)$$

Not surprisingly, it appears that one can use discrete FT spectrum to find the most probable sources of spectral artifacts.
In the first step, one shifts the centre of PSF (normalized to one at maximum) to the point where $\hat{S}(\vec{f})$ has a maximum absolute value $|I_{max}|$ (see also Figure 9). Then one subtracts a fraction $0 < \gamma \leq 1$ (called "loop gain") of the shifted PSF:

$$\hat{S}(\vec{f})^{(i+1)} = \hat{S}(\vec{f})^{(i)} - \gamma III(\vec{f} - \vec{f}_{max})$$
$$(40)$$

The extracted component gives rise to so-called "replica", G(f), which is hoped to reproduce the perfect spectrum S(f) at the end of the procedure:

$$\hat{G}(\vec{f}_{max})^{(i+1)} = \hat{G}(\vec{f}_{max})^{(i)} + \gamma \cdot I_{max}$$
$$(41)$$

Providing that the peak at the selected point was a real feature, one obtains a new spectrum $\hat{S}(\vec{f})^{(i+1)}$ with decreased level of artifacts.
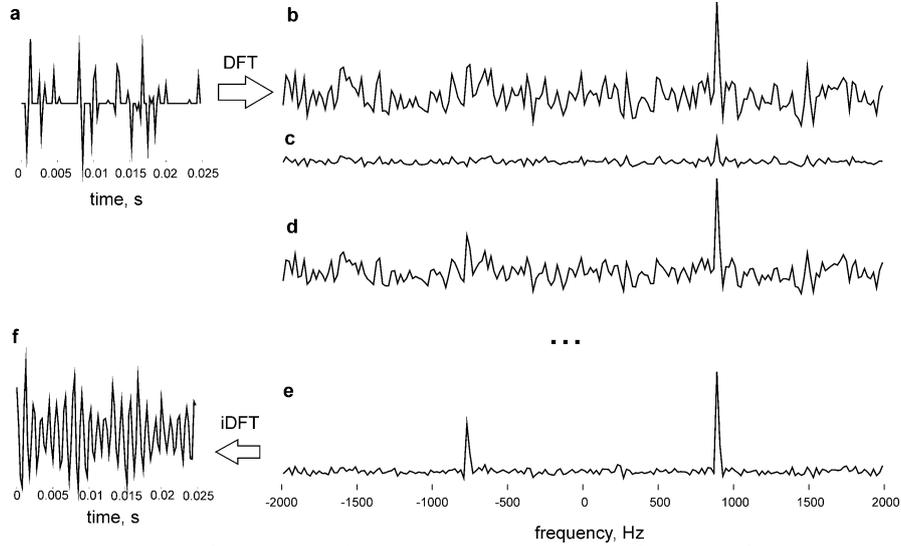
Fig. 9. The principle of CLEAN algorithm visualized on a simulation of 3 signals of relative amplitudes 1:5:10 and equal decay rates. The sparsely sampled signal (a) is Fourier transformed (b), then the mask (c) is subtracted to yield residual spectrum. Reconstruction after the first iteration is shown (d). The final result of CLEAN procedure (e) can be used to obtain reconstruction of time-domain signal (f).

In the next iteration, one can repeat the steps of (i) finding the most intense spectral amplitude and (ii) subtraction of shifted PSF. The whole procedure should be continued until there are no significant peaks in the spectrum. This condition can be formulated as follows:

$$\left| I_{max} \right| < \alpha \cdot \sigma_i$$

(42)

where σ_i is the estimated noise level in the i-th iteration, and α is usually a small integral value (3-5). It should be noted, that σ_i is a measure of both remaining artifacts and usual thermal noise $\varepsilon(\vec{f})$. Finally, the residual spectrum may be added to "replica" in order to retain smaller features that might have been omitted by the CLEAN algorithms, or to reintroduce usual noise $\varepsilon(\vec{f})$. The latter might be useful to judge which peaks selected during the iterations are false [85]. It was emphasized that displaying the "replica" without the addition of residual spectrum is merely a "cosmetic" operation and does not improve the sensitivity at all [86].

It is noteworthy that the uncertainty of peak amplitudes caused by

the presence of noise $\varepsilon(\vec{f})$ limits the capability of CLEAN algorithm to improve the quality of spectrum [87]. This, however, should apply for the most of reconstruction algorithms, e.g. similar conclusions were drawn for the maximum entropy method [88].

7.1.3 Discussion of the parameters of CLEAN

Apparently, CLEAN has two parameters which can affect both efficiency (in terms of computational effort) and accuracy of the final results. *Loop gain*, γ, determines how fast are the artifacts removed from the spectra in each iteration. Generally, it should reflect the probability that the selected peak is true (not an artifact nor a noise peak) and that the intensity observed in the spectrum I_{max} comes entirely from the component centered at $\vec{f}_{max}$. Therefore, small values should be used for spectra containing overlapped peaks [89], peaks broader than PSF [85, 89] or noisy ones. Alternatively, *loop gain* can take a variable value depending on the ratio $|I_{max}|/\sigma_i$.

Clearly, only the infinitesimally small value of loop gain ensures maximum safety of the procedure [84]. However, decreasing the loop gain causes a serious efficiency penalty, and the compromised values between 0.25 and 0.5 are typically used [86]. It has been pointed out that larger values can result in false splittings, especially when there is a mismatch of the "mask" line widths and the experimental ones [89].

The second parameter is the intensity threshold, which can be determined by the operator in advance and kept fixed, or evaluated dynamically on the basis of the current noise level in the spectrum. The former option requires a prior knowledge of noise amplitude, whereas the latter needs a robust method of measuring of the noise level.

Regarding the termination criterion, one should comment that there is a trade-off between the safety of peak identification and completeness of artifact suppression [89]. For example, the threshold of $5\sigma_i$ gives a great confidence that only genuine peaks are extracted, however, it also limits the benefits of the CLEAN procedure as the artifacts originating from less intense components remain in the spectrum.

It has been also agreed that a fixed number of iterations is difficult to apply in practice for NMR spectra and could lead to misinterpretations of the results of CLEAN algorithm [87]. Therefore, one should rather use the intensity threshold as the stopping condition.

Other authors [86] also noted that it is advantageous to use fine digitization in the frequency domain as it enables to precisely position the "mask" (PSF). On the other hand, this does not seem critical for the results and may unnecessarily increase the computational burden.

7.2 Development of the CLEAN algorithm

As noticed by Coggins and Zhou [81], if CLEAN is employed to suppress artifacts originating from irregular sampling, the artifact level varies greatly in the multidimensional spectrum along directly detected dimension. Consequently, it is impractical to use a fixed intensity threshold in this case. Apart from the commonly used dynamic threshold of $5\sigma_i$, it was suggested to employ the noise stabilization criterion, which stops the iteration if CLEAN does no longer efficiently remove artifacts. The condition was quantified as follows:

$$\overline{\sigma}_j \le (1+\tau)\overline{\sigma}_i \quad for \quad i-25 \le j < i \tag{43}$$

where $\overline{\sigma}_j$ denotes the average noise level measured in the *j*-th iteration. The tolerance for noise stabilization τ of approximately 0.05 was suggested. One may consider this condition a practical optimization of CLEAN algorithm in terms of numerical efficiency, not necessarily improving the quality of the final spectra.

In contrast to other implementations, Coggins and Zhou used the mask computed directly from the sampling function, without the knowledge of minimal signal line width. It was argued that such approach is more general, as broad or overlapped peaks can be represented sufficiently accurate by a superposition of narrow peaks. Indeed, frequently the resolution of spectra of biomolecules is

mostly determined by signal truncation, and the natural peak line widths can be neglected when using CLEAN in these applications. A different approach to CLEAN processing was suggested by Kazimierczuk and co-workers [60]. In their implementation, peaks are manually fitted in the initial spectrum using assumed shapes (e.g. Lorentzian or Gaussian, depending on the decay of sampling density employed). The list of peaks and their line widths, which can be considered the analytic equivalent of "replica", is then provided to the processing program. In the following, the artifacts generated by the peaks in the list are computed and subtracted from the initial spectrum, except for the peak positions and their vicinities. The procedure can be repeated if advantageous, e.g. if a significant number of (new) medium and small peaks were found after subtraction of the artifacts.

It should be noted, that the use of an appropriate analytic function instead of discrete "replica" may be beneficial as this (i) is less influenced by noise and (ii) allows to reproduce the "wings" of the resonances, which are neglected in the original CLEAN algorithm due to intensity threshold. On the other hand, if the line widths in the Fourier domain are mainly due to signal truncation, the fitted parameters poorly reflect the true signal properties, and this may affect the performance of described procedure.

A remedy for this was proposed by Stanek and Koźmiński [83]. In their significantly modified version, referred to as Signal Separation Algorithm (SSA), peaks are automatically found and fitted using the mono-exponentially decaying functions in the time domain. As a consequence, the simulated line shapes in the Fourier domain are affected by sampling process in the same manner as the real peaks in the spectrum. The advantages of this approach over the original CLEAN were demonstrated [83] (see also Figure 10). Another modification proposed by these authors regards the case of overlapped peaks or when decay parameters cannot be reliably established. It was suggested to find a replica that reproduces the observed peak shape in the iterative process. The idea to vary the amplitudes in replica until the desired peak shape is obtained clearly alleviates the problem of the appropriate value of *loop gain*.
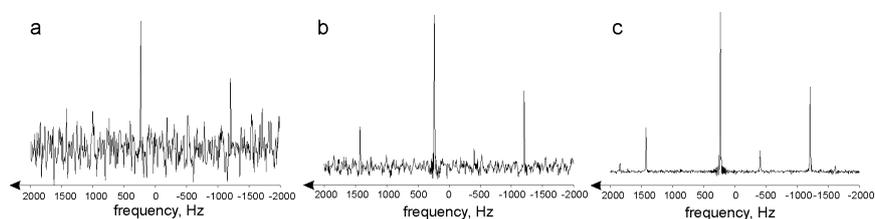
Fig. 10. Comparison of the efficiency of CLEAN (b) and SSA (c) shown on a simulated signal containing 6 components of relative amplitudes 1:2:4:8:16:32 and equal decay rates of 20 s^{-1}. Additionally, white Gaussian noise of σ=0.02 was present. Both algorithms started from the same initial nuFT spectrum (a), and the same threshold for peak detection equal to 5σ$_i$ was used. Spectral width of 4kHz, and max. evolution time of 70ms were set. 70 out of 280 points were sampled, yielding relative sampling density of 0.25.

7.3 The early applications of CLEAN to NMR spectroscopy

The original CLEAN algorithm was indented to effectively deconvolve the Fourier spectrum from the PSF. In the radioastronomy it was either impossible or impractical to arrange detectors on a regularly spaced grid due to malfunctioning of the part of equipment, occultations caused by the Moon or if telescopes were operating on a large area to provide high resolution maps [84]. The aim of CLEAN was to convert the map obtained from irregular and/or coarse grid of interferometers to that which would be obtained from the fine and complete grid.

As pointed out by Davies and co-workers [89] many high-resolution 2D spectra are not sparse and suffer rather from signal overlap or line shape distortions caused by several reasons (twisted-shape, truncation artifacts, inhomogeneous broadening). Although these difficulties seem quite different from those in radioastronomy, it became possible to employ essentially the same algorithm to alleviate these problems.

The idea was to construct the "mask" similar to undesired shape observed in the spectra, and use CLEAN to replace distorted peaks with those of a perfect Lorentzian shape. Shaka and co-workers [86] showed that the algorithm is capable to convert twisted-shape to double-absorbtion in 2D phase-sensitive *J* spectra of complex organic molecules. This was achieved by (i) locating the twisted-

shape peaks, (ii) simulating double-dispersion signals of the same line width at the same frequency coordinates, and (iii) subtraction of the latter from the original spectrum. Effectively, the most intense peaks were in double-absorbtion while those ignored by CLEAN remained in twisted-shape.

Similar approach was presented by Keeler in application to heteronuclear *J* spectra with highly truncated echo modulation [85]. Truncation of signal, used for sensitivity reasons, results in "sinc wiggles". These artifacts can be suppressed by apodization, however, at the expense of resolution. Keeler showed that CLEAN is an inexpensive alternative to maximum entropy method, which also can remove truncation artifacts without degrading resolution.

The difficulty that has arisen in both applications was to adjust the line width of the mask, in order to fit all signals. It has been suggested that if there is a mismatch of the line widths between the mask and experimental line shapes, one has to decrease *loop gain* and represent broad peaks as a superposition.

Davies and co-workers alleviated the problem of the optimal mask, by using experimental line shape of a well separated singlet resonance [89]. This was showed to enhance resolution of spectra more effectively than when simulated Lorentzian mask is employed. Additionally, CLEAN was compared to maximum entropy method, giving similar results in considerably shorter computational time. One should note, that the use of experimental aperture shape to compensate for spatial inhomogeneity of magnetic field is limited to the cases of high S/N, otherwise the mask is heavily biased by noise.

In all cases described above the fixed threshold (of a few per cent of the tallest peak) was used to terminate processing. This was possible as the noise level does not significantly vary during the iterations. The latter does not hold in the case of irregular sampling and more careful termination criteria have to be applied when deconvolving the PSF [87]. It has been showed on both experimental data and simulations that similar results can be obtained by CLEAN and maximum entropy method, and that CLEAN performs much better in recovery of missing samples than in extrapolation of a truncated signal [87]. As mentioned above, the success of CLEAN was limited by S/N.
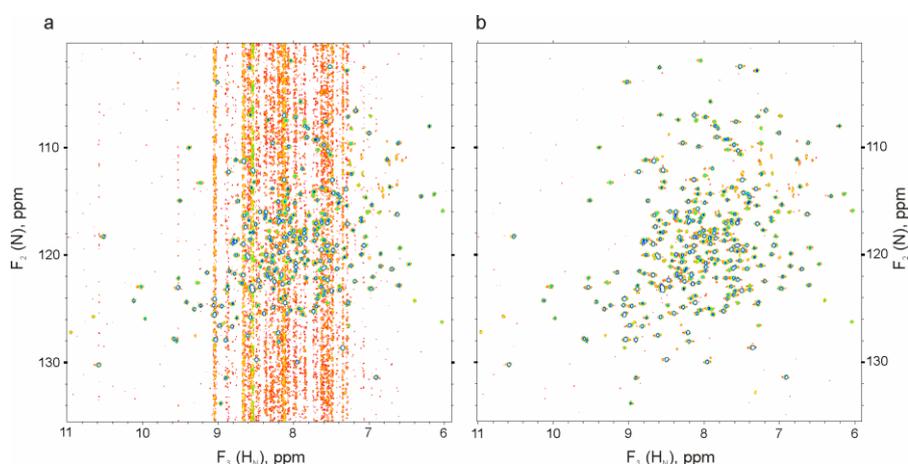
Fig. 11. F_2/F_3 projections (along $F_1(C')$ dimension) of 3D HNCO-TROSY spectra for maltose binding protein (371a.a., uniformly-deuterated, 0.5mM D_2O/H_2O 1:19 solution), obtained using sparse on-grid sampling, nuFT (a), and SSA processing (b). The data were recorded at the Varian 700 MHz spectrometer, assuming the spectral widths of 2.8 and 2.5 kHz in $F_1(C')$ and $F_2(N)$ dimensions, respectively. 1750 sampling points were generated using decaying sampling density ($\exp(-t^2/2\sigma^2)$, $\sigma=0.5$). Maximum evolution times of 30 and 50 ms were set, yielding relative density of $\theta=16.7\%$.

According to these observations, the power of CLEAN algorithm was utilized in high-dimensional (3D and 4D) NMR spectroscopy of proteins (see Figure 11), where *sparse sampling* has to be employed due to practical limitation on experiment time (see Section 9.4). In conclusion, the application of CLEAN algorithm to sparsely sampled data is especially beneficial if (i) the technique features good thermal sensitivity and (ii) a high dynamic range of peak amplitudes is expected. Otherwise, artifact suppression is hampered or irrelevant in view of the general noise level.

7.4 Algorithms related to CLEAN

It is noteworthy to mention that the principle of CLEAN algorithm was also utilized in several other processing methods [40, 70]. Kupče and Freeman adapted the processing scheme to remove ridges and false peaks present in the projection-reconstruction of 3D spectra [40]. As noted by the authors, it can be confidently assumed

that the tallest peak in the reconstruction is genuine. It is then possible to extract it *from the projections* and reconstruct the full spectrum again. As usual, the process can be repeated to further suppress projection-reconstruction artifacts until no significant peaks are present. At the final stage, the extracted peaks are reintroduced to the full spectrum.

Hyberts and co-workers described a "distillation" procedure which improves the quality of Forward Maximum entropy (FM) and l_1-norm reconstructions [70]. The purpose of this processing scheme is to divide FID signal into two components, one containing "tall" and another "small" spectral information. The division is performed in the Fourier domain according to the relative amplitude of each pixel to the most intense one, and both parts "small" and "tall" are inversely transformed to the time domain. The advantage is that FM performs the reconstruction on a sub-spectra of decreased dynamic range of peak amplitudes. This was shown to improve both linearity of the method and suppression of sampling artifacts. This is in analogy to CLEAN processing, where time domain signal is effectively split to the contributions from strong and weak signals. It is noteworthy, that the "distillation" procedure does not require any parameters and usually up to 8 iterations are sufficient.

8 FT as a tool for large evolution time domain

8.1 Features of "sampling noise"

As mentioned above, the nuFT does not find the optimal solution that fits to the experimental data. Spectra obtained by nuFT suffer from additional artifacts, which, in the case of random sampling, take a noise-like form. Luckily, they reveal also similar properties as thermal noise, i.e. artifact level is proportional to $\sqrt{N}$ (see Figure 12) and does not depend on a dimensionality of a signal, maximum evolution times nor spectral widths [45]. This fact may be proved in various ways (two of them were presented in [45]), below we will

present a new, simpler proof, based on known properties of Monte Carlo integration [66].

Discrete multidimensional Fourier transform of randomly sampled signal may be considered as an estimation of continuous multidimensional integral (Eq. 12) with Monte Carlo procedure. According to properties of Monte Carlo integration, associated with the law of large numbers and the central limit theorem, the approximate value of an integral with some finite integration volume V:

$$I = \int_V f(\bar{x})d\bar{x}$$

(44)

is given by:

$$Q = \frac{V}{N}\sum_{i=1}^{N} f(\bar{x}_i)$$

(45)

with an expected value equal to the value of continuous integral (unbiased estimator):

$$E[Q] = I$$

(46)

and the variance decreasing with N:

$$\mathrm{var}(Q) = V^2 \frac{\mathrm{var}(f(\bar{x}))}{N}$$

(47)

Thus, the error of approximation is decreasing with $\sqrt{N}$. The only difference between signal processing and numerical integration is the way how samples are obtained. Instead of calculating the values of function at randomly selected points, as it is done in the Monte Carlo procedure, the integrated function is *experimentally measured* at these points (or, more strictly, measured and multiplied by the transform kernel). Nevertheless, the way how points are obtained does not affect general conclusions, i.e.: that the estimator is unbiased and converges to the perfect, artifact-free spectrum with

growing number of sampling points and that the relative error of the result (S/A ratio) is inversely proportional to $\sqrt{N}$. Notably, the estimation error does not depend on parameters that cause sampling-related problems (i.e. limited resolution) in conventional approach e.g. dimensionality of a signal and maximum evolution time (see Figure 13). This feature makes random sampling a perfect tool for high-dimensional (4D,5D,6D etc.) NMR experiments [45] with quite high absolute number of sampling points (not necessarily meaning high sampling density!). For the same reasons, Monte Carlo is known to be a favorable method for integration of high-dimensional functions [66]. It is also noteworthy, that other features of random sampling processed with nuFT have their equivalents in Monte Carlo method. For instance, stratified sampling is known to reduce variance of an integral estimation [66].
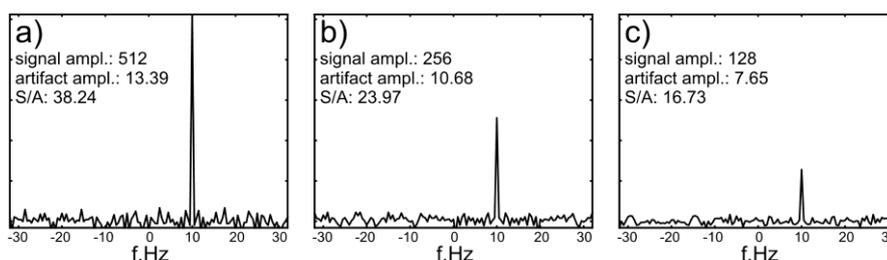
Fig. 12. Peak amplitude, artifact level and signal-to-artifact ratio for spectrum of non-decaying signal of frequency 10 Hz, sampled with: a) 512, b) 256, c) 128 points. Uniform random sampling was used.

Besides the absolute number of points, the artifact level is also inherently associated with a number of peaks and their intensities (as artifacts are "part" of Point Spread Function, see Figure 14). Thus, the more peaks in a spectrum, the lower is the average signal-to-artifact ratio. This makes nuFT processing more challenging when applied to spectra featuring large number of signals with high dynamic range of peak intensities (e.g. NOESY). In this case artifact-cleaning algorithms may be employed (see Section 7).
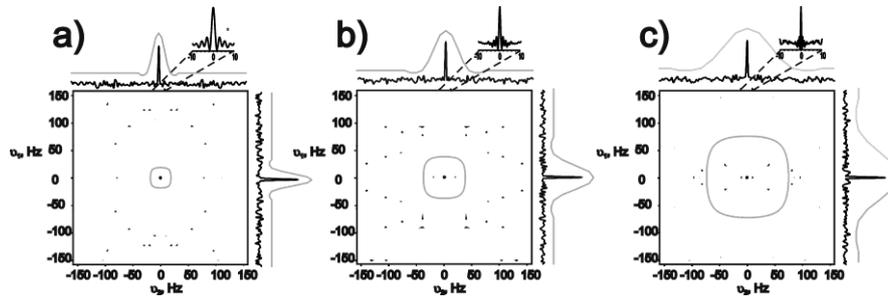
Fig. 13. 2D cross-sections from simulated: (a) 3D, (b) 4D, (c) 5D spectra. The threshold was set at 10% of peak intensity; 256 time points were generated randomly with uniform distribution and maximum evolution time of 0.4 s (panel a), 0.8 s (panel b) and 1.6 s (panel c), in all dimensions. The distance between spectral points was set to the reciprocal of maximum evolution time in order to hide the effect of signal truncation. The insets showing a spectral line narrowing obtained by MFT using higher digital resolution. Simulation was repeated for the conventional set of 256 points, with the Nyquist rate of 16 × 16 (panel a), 8 × 8 × 4 (panel b), and 4 × 4 × 4 × 4 (panel c). Peaks obtained in such way are shown with grey line. Reprinted with permission from Ref. [78].
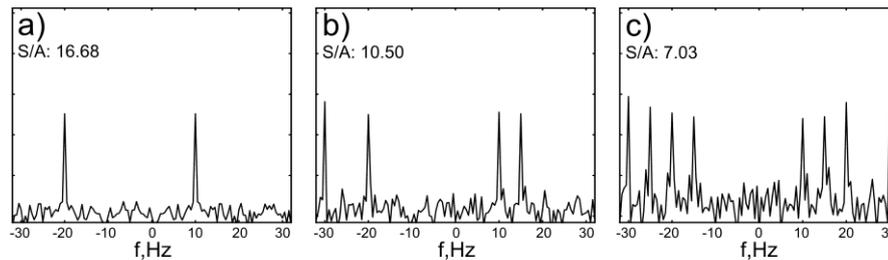
Fig. 14. Decrease in signal-to-artifact ratio with growing number of peaks: a) 2, b) 4, c) 8. S/N was calculated using "true" peak amplitude (not influenced by artifacts).

8.2 Sparse MFT (SMFT)

According to Eq. (12) it is possible to arbitrarily choose frequency points for FT, e.g. to calculate just an interesting region(s) of a spectrum. This approach is of a particular use when dimensionality and/or resolution is high and the full spectral matrix would be of an extremely large size. There are various possibilities for restricting spectral space to the regions of interest, depending on the type of spectrum and type of information to be extracted [78]. All of them base on prior examination of other (simpler) spectra. The restriction not only allows to save disk space, but also accelerates calculations

and facilitates data analysis.

8.2.1 "Slice" MFT

In spectra of high dimensionality peak coordinates in some of spectral dimensions are usually known from the spectrum of lower dimensionality (later called "basic spectrum"). The complete and regular frequency grid is not needed in these dimensions and they may be reduced to a set of frequencies corresponding to the tops of peaks [79] (see Figure 15). The number of lower-dimensional (e.g. 2D) cross-sections obtained with this approach is equal to the number of peaks found in the basic spectrum. Noteworthy, the basic spectrum, used for frequency selection, should be also recorded with high resolution, as an accuracy in determination of peaks frequencies is crucial here. Such a procedure dramatically reduces the amount of data to be stored. In general, the size of data matrix of N-dimensional spectrum is equal to:

$$size = m_1 \cdot m_2 \cdot ... \cdot m_N$$

(48)

where: m_i is a number of spectral points in *i-th* dimension.
If the frequencies of first k dimensions are "reduced" during FT, the data matrix size becomes:

$$size = ns \cdot m_{k+1} \cdot m_{k+2} \cdot ... \cdot m_N$$

(49)

where: *ns* is a number of frequency sets obtained from a lower-dimensional spectrum used for SMFT. For instance, let us assume that number of spectral points in each dimension of 5D data set is equal to 128 and SMFT is performed on the basis of a 3D spectrum containing 150 peaks. In this case, the total size of resulting set of 2D planes will be $128 \cdot 128 \cdot 128 / 150 \cong 13981$ times smaller than the size of full 5D spectrum, which in practice means reduction of the file size from about 100 GB to approximately 10 MB. Moreover, a set of lower-dimensional spectra is easier to handle than one spectrum of high dimensionality (see Section 9).
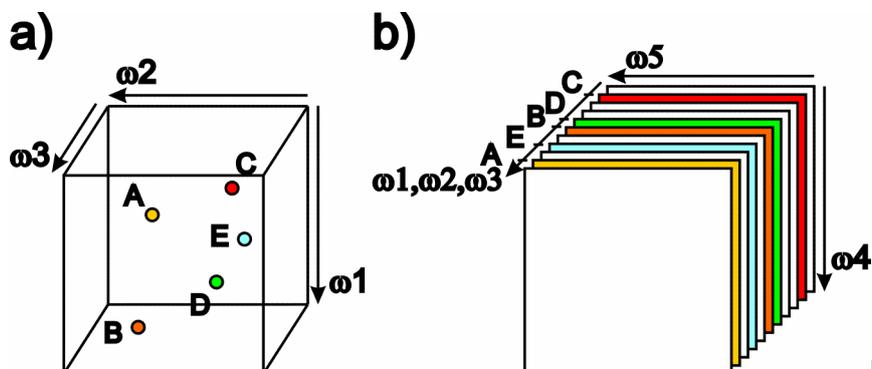
Fig. 15. The idea of a "slice" SMFT. a) A scheme of a 3D spectrum. Frequency coordinates of peaks from this spectrum (labeled A-E) are used as the basis for SMFT calculation. b) A scheme of a 5D spectrum. Three frequency dimensions ω_1, ω_2 and ω_3, which correspond to nuclei observed in 3D spectrum are symbolized by one axis, two other dimensions (ω_4 and ω_5) are shown on separate axes. Only 2D (ω_4–ω_5) cross-sections that contain peaks (marked with colors) are calculated in SMFT. Reprinted with permission from Ref. [78].

8.2.2 "Cube" MFT

Using techniques of extraordinary resolution, it is possible to efficiently measure peak splitting (E.COSY pattern) associated with internuclear couplings [77]. By increasing maximum evolution times one can reach peak width determined practically only by relaxation rate. However, ultra-narrow peaks require enhanced digital resolution (number of points per Hz) to be properly visualized. This often causes the need to use another procedure employing reduced frequency space.

Prior to processing of such a high-resolution data, positions of peaks should be roughly determined from an equivalent decoupled (i.e. with singlets) spectrum (or spectra) of lower resolution. Afterwards, the spectrum of high resolution is calculated only in a close vicinity of these peaks positions, resulting in a set of full-dimensional "cubes" (see Figure 16). In each "cube" the numerical resolution should be sufficiently high to visualize the multiplets and determine coupling constants. Again, reduction of the required disk space is significant. The size of data matrix is reduced from that defined by Eq. (48) to the following value:

$$size = ns \cdot m_1 \cdot \frac{sw_1^{loc}}{sw_1} \cdot m_2 \cdot \frac{sw_2^{loc}}{sw_2} \cdot ... \cdot m_N \cdot \frac{sw_N^{loc}}{sw_N} \qquad (50)$$

where: ns is a number of frequency sets used for SMFT, sw_i is spectral width in dimension i of full spectrum, and sw_i^{loc} is spectral width in dimension i of a single "cube".

For example, in case of 4D spectrum, when spectral width of a "cube" is in each dimension 10 times smaller than full spectral width in this dimension, and number of "cubes" is 150, the data set is reduced about 40000 times. Typically, it can result in reduction of disk space requirement from tens of TB to the order of GB.
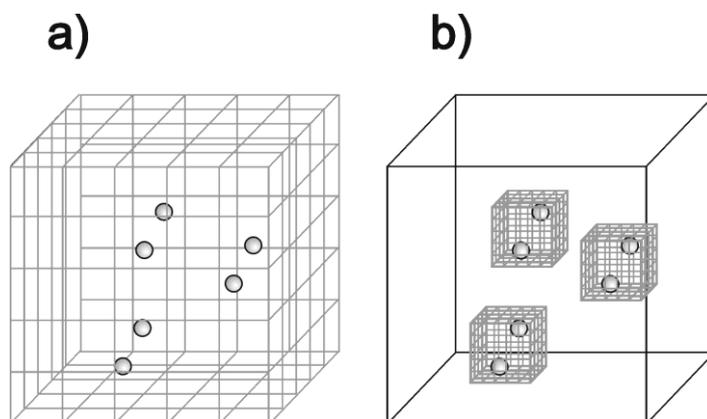
Fig. 16. The idea of a "cube" SMFT. a) A scheme of a full 3D spectrum, containing peaks revealing E.COSY multiplet structure. The digital resolution is too low to properly approximate narrow components of multiplets. b) A scheme of a set of "cubes", calculated just in vicinities of peaks, featuring much higher digital resolution. Determination of small coupling constants is possible. Reprinted with permission from Ref. [78].

9 Applications

The interpretation of one- and two-dimensional spectra of large biomolecules such as proteins and nucleic acids is usually impossible due to a large number of highly degenerated peaks. Hence, even for the medium-sized molecules, it is necessary to use

isotopic enrichment with ^{13}C and ^{15}N nuclei, and to perform triple-resonance 3D NMR experiments for resonance assignment and extraction of structural constrains. However, as we pointed out above, the resolution of conventionally acquired 3D spectra, is limited by sampling requirements. Therefore, it is rarely possible to obtain line widths close to the natural ones in a reasonable time, even for very fast-relaxing molecules. The conventional 4D spectra, as for example ^{15}N,^{13}C or ^{13}C, ^{13}C-edited NOESY experiments, are rarely employed owing to the low evolution times achievable. On the other hand, NMR spectra of biomolecules feature relatively narrow and well-defined spectral regions as for example H_N, N_H, C', Cα and CαCβ in proteins. This feature allows the development of numerous multidimensional experiments, which correlate spin interactions in different dimensions. Thus, the most important applications of sparse sampling techniques are focused on the important field of structural studies of biomolecules in solution. Sparse non-uniform sampling and Fourier transform enable acquisition and processing of multidimensional NMR spectra featuring extraordinary resolution, as for example 4-6D NMR spectra dedicated to resonance assignment, techniques for precise determination of coupling constants from 3-4D experiments and proton-proton contacts from well-resolved NOESY spectra.

9.1 Development and implementations

The early applications of sparse sampling and FT processing were devoted rather to demonstrate the features of proposed methods than for cases of a really demanding nature. Kazimierczuk and coworkers compared 3D HNCO spectra of human ubiquitin employing radial and spiral sampling, showing significant advantages of the latter [44]. Shortly after that, Marion demonstrated FT processing in polar coordinates in application to radially sampled 3D HNCO of human ubiquitin [55]. The next was the work of Coggins and Zhou [56], who formulated the expression for polar FT and applied it for 3D TROSY-HNCO for ^{13}C/^{15}N/^{2}H-labeled OTU protein. In the consecutive works all three groups concentrated on the reduction of the artifact level . The Koźmiński's and Marion's groups switched

to random sampling, motivating it by the lower intensity and noise-like nature of artifacts in randomly sampled spectra. In the following works, the issue of approximation of Fourier integral was discussed. Kazimierczuk and co-workers [46] demonstrated that the surface integration using Delaunay triangulation improves S/A only for sampling above the Nyquist density. It was also shown that in the case of unweighted FT the S/A ratio does not depend on the relative samples density (see Figure 8). The usability of the method was verified on 3D HNCA, HNCACB and ^{15}N-edited NOESY experiments on ubiquitin, using random sampling with exponential and Gaussian distributions of sampling points. Later on, the same authors [60] showed that the simple regularization of the samples distribution reduces artifacts in the signal vicinity. Moreover, it was demonstrated that this effect is more pronounced in comparison with Voronoi tessellation used as an integral quadrature rule. Additionally, in this work the usability of a simple variant of CLEAN algorithm (see Section 7) was demonstrated on 3D ^{15}N-edited NOESY spectrum of ubiquitin. Pannetier and coworkers [64], for the first time applied random sampling and FT processing for intrinsically unstructured protein, namely 60-residue N_{TAIL} (443–501) fragment of nucleoprotein N from the paramyxovirus Sendai. They obtained backbone resonance assignment using two 3D CBCANH and CBCA(CO)NH experiments with 6.5-fold undersampling. In the same time, Coggins and Zhou introduced concentric ring sampling, demonstrating its advantages over initially used radial alternative, and employed it for 3D HNCO of uniformly ^{13}C, ^{15}N labeled spectrum of the B1 domain of protein G (GB1) [59]. The next development in Zhou's group was concentric shell sampling adjusted to a fine grid which was employed for the 4-D HCCH-TOCSY [81]. In this work artifact suppression was accomplished by an adaptation of CLEAN algorithm (see Section 7).
The influence of different constrained random sampling schedules on Point Spread Function was further investigated by Kazimierczuk and coworkers considering both artifact level and distribution [61]. It was shown that Poisson disk sampling provides the largest low-artifact area in the signal vicinity. The new sampling schemes were verified by application to the 3D HNCACB and ^{15}N-edited NOESY-HSQC acquired for human ubiquitin. The analysis of signal-to-artifact ratio with respect to relative sampling density and

dimensionality was analyzed in the next work from the same group [45]. It was proven that for random sampling S/A ratio depends neither on sampling density nor dimensionality of the experiment. These results were experimentally confirmed by acquisition of 5D HC(CC-TOCSY)CONH, performed for doubly labeled human ubiquitin within 0.0054 % of time necessary for analogical conventional experiment.

9.2 Easy resonance assignment in proteins using the spectra of high dimensionality

High resolution and dimensionality achievable in spectra acquired with the use of sparse random sampling and processed by FT feature a significant improvement in peak dispersion. This facilitates resonance frequency assignment especially in demanding cases such as, intrinsically disordered proteins. The first such example, mentioned above, was the backbone assignment of intrinsically unstructured 60-residue N_{TAIL} (443–501) fragment of nucleoprotein N from the paramyxovirus Sendai using the 3D experiments [64]. After the feasibility of 5D experiments acquired by random sampling and SMFT processing by Kazimierczuk et al. was demonstrated [45], the same group proposed a set of 4D (HNCOCA, HNCACO, HNCACACB, HN(CA)NH and HabCabNH) [79], and later 5D (HN(CA)CONH, HabCabCONH) experiments [78] dedicated to the effective protein backbone signal assignment. All of these techniques employ sparse random sampling and FT processing to achieve high resolution spectra in a tiny fraction of time needed conventionally. The 4D experiments were tested on two proteins differing in size, i.e. a protein interacting with NIMA-kinase from Cenarcheaum symbiosum (96 a.a. residues) and maltose binding protein (371 a.a. residues) (see for example Figure 17).
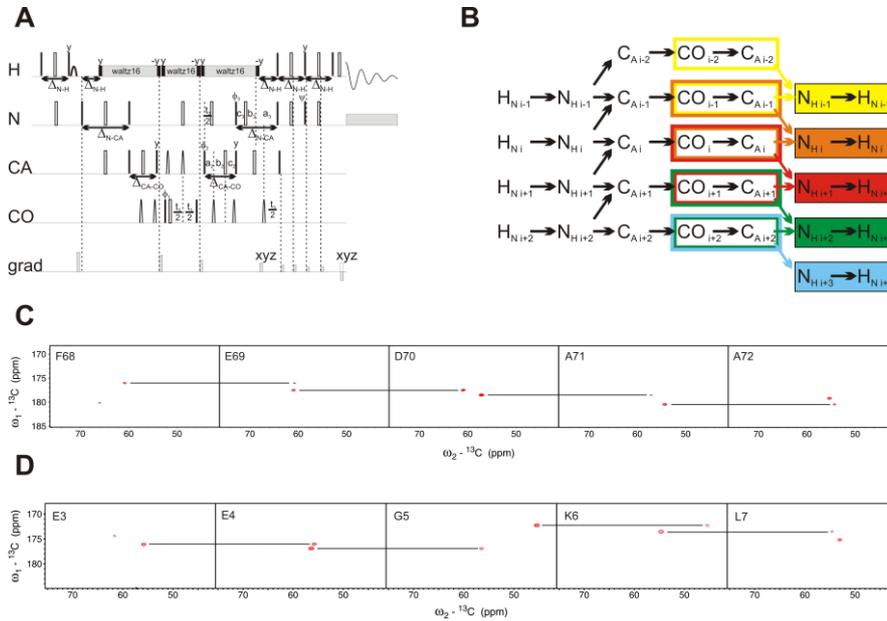
Fig. 17. Example of application of 4D HNCACO technique: (Panel A) Pulse sequence, evolution for CO is in the real-time mode, and for N and CA in semi-constant-time mode ($a_i = (t_i + \Delta)/2$, $b_i = t_i(1 - \Delta/t_{maxi})/2$, $c_i = \Delta(1 - t_i/t_{maxi})/2$) or constant-time mode ($a_i = (\Delta + t_i)/2$, $b_i = 0$, $c_i = (\Delta - t_i)/2$), where Δ stands for $\Delta_{N–CA}$ and $\Delta_{CA–CO}$, respectively, t_i is the evolution time in i-th dimension and t_{maxi} is the maximal length of evolution time delay. Delays were set as follows: $\Delta_{N–H}$ = 5.4 ms $\Delta_{N–CA}$ = 22 ms $\Delta_{CA–CO}$ = 6.8 ms. (Panel B) Coherence transfer in the peptide chain. Amide nitrogen and proton frequencies (filled colored rectangles) are fixed during Fourier transformation. Each plane contains CO–CA peak for i and i - 1 residue. (Panel C) 2D spectral planes for CsPin protein obtained by SMFT procedure performed on the 4D HNCACO randomly sampled signal (Poisson disk sampling) with ''fixed'' frequencies obtained from 3D HNCO peak list. (Panel D) 2D spectral planes for MBP obtained in the same manner. Reprinted with permission from Ref. [79].

The feasibility of the 5D techniques was demonstrated using the sample of 5–79 fragment of bovine Ca^{2+}-loaded Calbindin D9K P47M mutant [79]. However, the true test of the new assignment strategy was performed on the particularly demanding case of δ subunit of RNA polymerase from Bacillus subtilis containing a disordered C-terminal region of 81 amino acids with a highly repetitive sequence [90]. While the backbone assignment of this protein appeared to be unachievable using conventional 3D techniques, the strategy based on the new, 5D experiments (HN(CA)CONH, HabCabCONH and HC(CC-TOCSY)CONH), provided a complete backbone and side-chain assignment (see Figure 18).

Fig. 18. ^{1}H,^{15}N-HSQC spectra of RNA polymerase d subunit (A, B, C) Panel A entire spectrum and Panel B with the central region expanded. Central region of ^{1}H,^{15}N-HSQC spectrum (C) and the corresponding region of a 2D cross-section extracted from the 5D HN(CA)CONH spectrum (right). The 2D cross-section was obtained by fixing frequencies in ω_3, ω_4, and ω_5 dimensions to the values of chemical shifts of ^{13}C' of I157, ^{15}N of I158, and ^{1}H$_N$ of I158, respectively. Peaks corresponding to the sequential and intraresidual correlations are displayed in black and red, respectively. Experiment was acquired within 20 hrs on 700 MHz spectrometer using RT-probe. Only a fraction of 0.00034 % points was collected in indirect time domains. See Ref. [90] for further details.

9.3 Determination of coupling constants in proteins

Backbone scalar couplings are widely used in NMR studies of structure and dynamics of biomolecules [91]. Additionally, there is also a substantial interest in precise determination of residual dipolar couplings for structural studies of weakly oriented biomolecules. Most of the relevant coupling constants in proteins are rather small – of the magnitude from a few to hundred hertz. Therefore, in order to achieve the sufficient resolution in indirectly measured dimensions, the majority of traditional methods devoted to coupling constants determination in biomolecules are limited to two-dimensional techniques, which frequently suffer from peak overlap. However, the random sampling of evolution time domain allows one to obtain spectra of resolution that is limited only by transverse relaxation and suffices to differentiate multiplet components. Moreover, when couplings with passive spins are resolved in two or more dimensions, the E.COSY [92] multiplet patterns provide valuable information about relative signs of coupling constants. Kazimierczuk and co-workers [77] showed an example of a 3D HNCO-$C\alpha$-coupled spectrum of ubiquitin protein. Each peak in this spectrum reveals 3D E.COSY pattern due to couplings with two passive $C\alpha$ spins. Thus, six coupling constants of H_N, N and C' with intra- and inter-residual $C\alpha$ spins can be determined. The resolution achieved in this experiment would require over a month of conventional acquisition, making it impractical. The coupling constants measured from 3D HNCO-$C\alpha$-coupled experiment revealed correlation with φ and ψ protein backbone torsional angles. Later on, a feasibility of determination of a 4D E.COSY patterns was also shown and exemplified with the 4D HNCACO-{$H\alpha$} experiment for the sample of 5–79 fragment of bovine Ca^{2+}-loaded Calbindin D9K P47M mutant [78], see Figure 19. In this experiment the "cube"-SMFT procedure was employed in order to achieve extraordinary disc space savings.
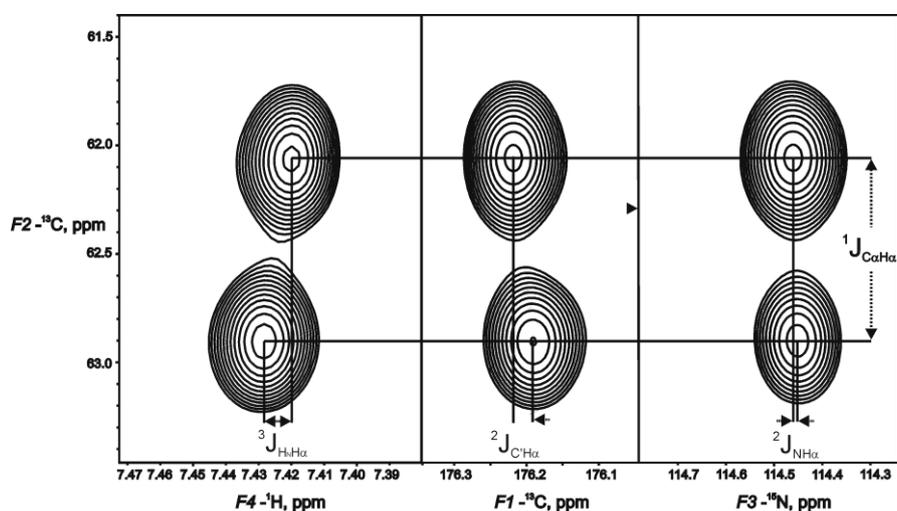
Fig. 19. The experimental example of ultra-high resolution multidimensional NMR spectra obtained by the proposed technique: I74 intra-residual resonance from 89-h 4D HNCACO-{H$_\alpha$}-coupled experiment acquired for 5-79 fragment of bovine Ca^{2+}-loaded Calbindin protein. Depicted cross-sections of 4D "cube" 50 × 450 × 40 × 100 Hz surrounding the peak allow determination of coupling constants from resolved 4D E.COSY pattern. $^1J_{C\alpha H\alpha}$ = 135.9 Hz, $^3J_{HNH\alpha}$ = 5.8 Hz, $^2J_{C'H\alpha}$ = -5.0 Hz, $^2J_{NHH\alpha}$ = -1.0 Hz with numerical resolution of 0.4 Hz/point, 1.7 Hz/point, 0.2 Hz/point and 0.7 Hz/point in dimensions F$_1$, F$_2$, F$_3$ and F$_4$, respectively. Reprinted with permission from Ref. [78].

9.4 Heteronuclear-edited NOESY experiments

NOESY experiments are still the primary source of structural information. The presence of the cross-peaks in NOESY spectra indicates spatial proximity of nuclei, and their integral is proportional to the r^{-6}, where r denotes internuclear distance. However, NOESY spectra are significantly more difficult to obtain in comparison with other NMR techniques. The most important differences are: a large number of correlation peaks, dependent on the number of interacting proton nuclei, and a high dynamic range of peak amplitudes up to two-three orders of magnitude. Consequently, NOESY spectra require an excellent sensitivity and almost perfect suppression of spectral artifacts. Moreover, in order to preserve relationship between peak integral and internuclear distances, the linearity of the method should be maintained. Thus, this type of applications is very demanding for all sparse NMR techniques. In

the case of nuFT processing effective artifact suppression is necessary.

Kazimierczuk and co-workers applied their semi-automatic CLEAN procedure to suppress artifacts in randomly sampled ^{15}N-labelled NOESY-HSQC spectrum of ubiquitin [60]. It was demonstrated that the process does not systematically influence relative peak amplitudes, and is therefore applicable to NOESY spectra. Similar conclusions were later drawn by Stanek and Koźmiński [83], and by Werner-Allen and co-workers [82], who compared their reconstructions with conventionally sampled three-dimensional spectra of the *same spectral resolution*. The algorithm proposed by Kazimierczuk and co-workers was later applied also to higher-dimensional experiments [78].

Coggins and Zhou implemented the CLEAN algorithm to process four-dimensional spectra [81], with only slight modifications with respect to the original procedure from radioastronomy. The advantages of CLEAN processing in conjunction with Randomized Concentric Shell Sampling were demonstrated on 4D HCCH-TOCSY spectrum of 56 a.a. GB1 protein. In this experiment, 1.2% of samples were used, and CLEAN was shown to decrease the apparent noise from 2.4 to 1.4 of thermal noise level on average. The same program was later used to process 4D amide-amide diagonal-suppressed TROSY-NOESY-TROSY (ds-TNT) spectrum of 23kDa C13S Sssu72 protein [82]. The largest decrease in apparent noise level due to CLEAN process was 22%. The application of sparse sampling and FFT-CLEAN processing allowed a more than 10-fold reduction in experimental time in comparison with the conventional approach to acquisition. The experiment was shown to provide valuable information on distance restraints between amide protons, by avoiding the ambiguities and frequent resonance overlap typical for 3D NOESY spectra of large proteins (see Figure 20).
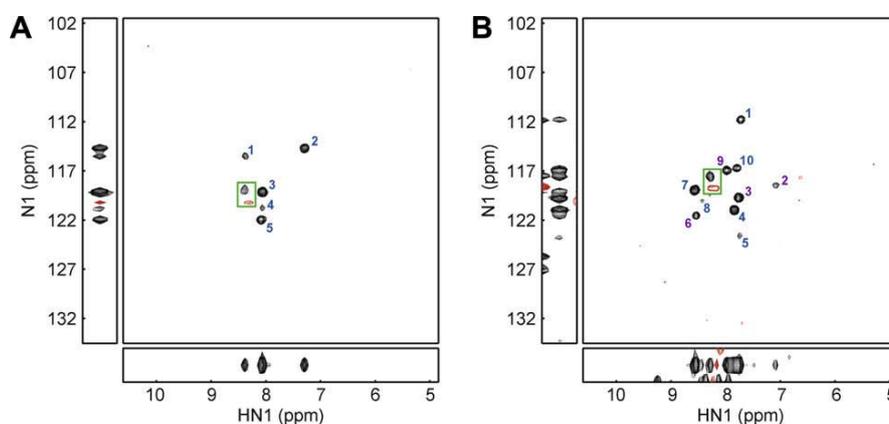
Fig. 20. F_1 (H_N) / F_2 (N) cross-sections from 4D amide-amide ds-TNT spectrum of C13S Ssu72 protein. Residual diagonal peaks of Ile176 (panel A) and overlapped Leu72 and Asn92 (panel B) are enclosed with green boxes. Corresponding strips in each panel, plotted from conventionally sampled 3D ds-TNT spectra, show severe overlap of the amide-amide cross-peaks. On the contrary, in the sparsely sampled 4D ds-TNT spectrum the peaks were clearly resolved and assigned. Reprinted with permission from Ref. [82].

A more challenging example was demonstrated by Stanek and Koźmiński [83], who applied their algorithm to 3D ^{15}N- and ^{13}C-labelled NOESY spectra of ubiquitin *without* suppression of diagonal peaks. The efficiency of artifact suppression was investigated by comparison of the reconstruction with the conventionally acquired reference spectrum. Less than 2% of peaks were missing, and about 1.5 % false peaks were reported. The correlation coefficient between peak volumes of R^2=0.998 was obtained.

9.5 3D spectra of complex organic compounds

Comparing to the progress and a variety of new multidimensional methods proposed in the area of biomolecules, in the field of organic molecules the development is slower, what is mainly caused by less demanding applications and additional experimental limitations. However, in the case of complex organic molecules it is sometimes necessary to add the third dimension to separate crowded, overlapping signals, in order to avoid ambiguities in the spectral assignment. The use of unconventional approaches to the acquisition of multidimensional NMR signal makes it possible to record 3D

NMR spectra of small molecules in shorter experimental time. So far there are only very few examples of using 3D NMR experiments dedicated to the spectral assignment [93-97] and the measurement of coupling constants [98].

Generally, in organic chemistry two-dimensional spectra are widely used, while 3D NMR spectra of small molecules were hardly achievable, because of very long measurement time required in the conventional approach. In contrast to proteins, organic compounds at the natural isotopic abundance are more demanding due to the low sensitivity and the necessity of sampling the wide frequency range especially in ^{13}C dimension. On the other hand the slower transverse relaxation rates allow to achieve narrow peaks, which again is limited by sampling. That is why in many cases these problems precluded the full assignment of NMR signals and the evaluation of coupling constants of organic compounds. Due to the employment of non-uniform sampling the application of multidimensional NMR spectra in the structure investigation of organic molecules became practically possible.

Recently, the method employing Dynamic Nuclear Polarization for the recording heteronuclear 2D NMR spectra of small drug-like molecules was proposed by Ludwig and coworkers [99]. This method ensures significant improvement in sensitivity due to the high spin polarization, but limits the number of points sampled in indirectly detected dimension, so the combination with non-uniform sampling scheme was necessary.

Three-dimensional NMR spectra based on random sampling of the evolution time space followed by MFT processing were successfully applied by Misiak and Koźmiński in the structural analysis of complex organic compounds [93]. A three new 3D NMR techniques (TOCSY-HSQC, COSY-HMBC and HSQMBC), which allow the spectral assignment have been proposed. The comparison of 3D spectra of strychnine recorded in the conventional way with that acquired using randomly distributed data points in the evolution time space revealed that by using this new approach it is possible to acquire 3D spectra in the reasonable experimental time, while retaining high resolution in indirectly detected domains (see Figure 21). The use of 3D TOCSY-HSQC and 3D COSY-HMBC allowed for the complete assignment of ^{1}H and ^{13}C chemical shifts of natural abundance prenol-10 [94], what was earlier impossible by

employing 1D and 2D spectra, mostly because of the signal overlapping caused by similarity of the ten isoprene units. The application of 3D HSQC-TOCSY spectra with E.COSY- type multiplets enabled the accurate determination of heteronuclear coupling constants of organic molecules in an overnight experiment [98].

In the case of natural abundance complex organic compounds playing important biological roles, the detailed structural analysis is very important. We believe that in the future recording of sparsely sampled 3D NMR spectra should become a routine procedure also for the structural analysis of complex organic molecules.
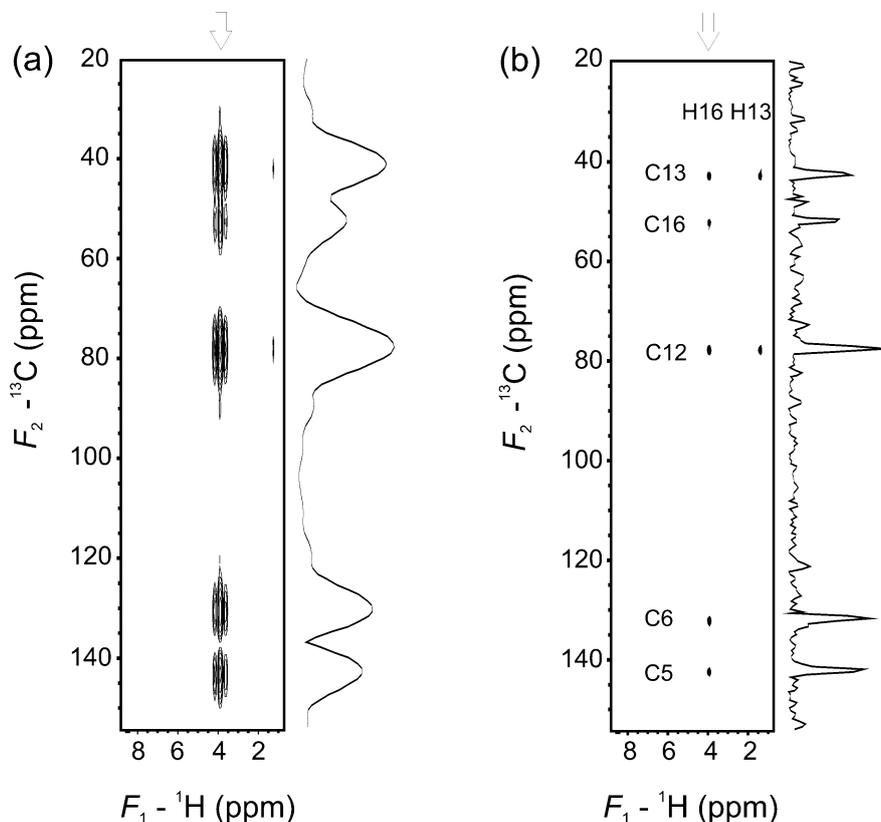
Fig. 21. Comparison of 3D COSY-HMBC F1/F2 cross-sections for F3 (1H) = 3.904 ppm, i.e. resonance frequency of the H16 atom of the strychnine molecule: (a) conventional and (b) random sampling of t1/t2 evolution time space. The spectra were recorded in the same experimental time, and transformer with the resolution of 128 x 256 x 1024 points in F1, F2 and F3, respectively. The vertical arrows indicate the positions of the extracted traces. Reprinted with permission from Ref. [93].

10 Conclusions

The application of sparse sampling for the acquisition of multidimensional NMR spectra causes the presence of spectral artifacts. They appear in a regular form (e.g. ridges, rings) for the regular sampling, and resemble noise for the case of random sampling. The spectral reconstruction aims to obtain spectrum with minimized artifact level. Among a variety of reconstruction methods the Fourier Transform has favorable computational requirements. The important feature of off-grid random sampling and FT processing is the independence of artifact intensity of the degree of sparseness, and decreasing of artifacts with the square root of number of sampled points. Therefore, it should not be applied for the acceleration of experiments attributing conventional resolution, which is a usual task of "fast NMR" techniques. FT is rather the method of choice for the acquisition and processing spectra of high dimensionality (4-6D) or of high resolution approaching natural line-width. Frequently, the artifact level in such spectra is low enough to allow their interpretation without further processing. However, for the analysis of high dynamic range spectra featuring a large number of signals, as for example NOESY experiments, additional artifact "cleaning" is required. Until now, the number of such applications is still minor, however, we expect that it will grow systematically in parallel with dissemination of the necessary software. We believe that the random sparse sampling and FT processing could be aimed for a variety of new applications, especially in the field of NMR-based structural studies of biomolecules.

Acknowledgments

References

1. Ernst RR, Anderson WA (1966) Application of Fourier Transform Spectroscopy to Magnetic Resonance. Rev Sci Instrum 37: 93-102
2. Jeener J, Lecture presented at Ampere International Summer School II, Basko Polje, Yugoslavia 1971
3. Aue WP, Bartholdi E, Ernst RR (1976) Two-Dimensional Spectroscopy. Application to Nuclear Magnetic Resonance. J Chem Phys 64: 2229-2246
4. Wüthrich K (1986) NMR of Proteins and Nucleic Acids. John Wiley and Sons, New York, NY
5. Montelione GT, Wagner G (1989) Accurate Measurements of Homonuclear H-N-H-Alpha Coupling-Constants in Polypeptides Using Heteronuclear 2D NMR Experiments. J Am Chem Soc 111: 5474-5475
6. Ikura M, Kay LE, Bax A (1990) A Novel-Approach for Sequential Assignment of H-1, C-13, and N-15 Spectra of Larger Proteins - Heteronuclear Triple-Resonance 3-Dimensional NMR-Spectroscopy - Application to Calmodulin. Biochemistry 29: 4659-4667
7. Bax A, Grzesiek S (1993) Methodological Advances in Protein NMR. Acc Chem Res 26: 131-138
8. Yamazaki T, Lee W, Arrowsmith CH, Muhandiram DR, Kay LE (1994) A Suite of Triple Resonance NMR Experiments for the Backbone Assignment of 15N, 13C, 2H Labeled Proteins with High Sensitivity. J Am Chem Soc 116: 11655–11666
9. Sattler M, Schleucher J, Griesinger C (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Prog Nucl Magn Reson Spectrosc 34: 93-158
10. Moskau D (2002) Application of real time digital filters in NMR spectroscopy. Concepts Magn Reson 15: 164-176
11. Szantay C (2008) NMR and the uncertainty principle: How to and how not to interpret homogeneous line broadening and pulse nonselectivity. IV. (Un?)certainty. Concepts Magn Reson A 32A: 373-404
12. Schanda P, Kupče Ē, Brutscher B (2005) SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. J Biomol NMR 33: 199-211
13. Schanda P, Van Melckebeke H, Brutscher B (2006) Speeding up three-dimensional protein NMR experiments to a few minutes. J Am Chem Soc 128: 9042-9043
14. Lescop E, Schanda P, Brutscher B (2007) A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. J Magn Reson 187: 163-169
15. Frydman L, Scherf T, Lupulescu A (2002) The acquisition of multidimensional NMR spectra within a single scan. Proc Natl Acad Sci U S A 99: 15858-15862
16. Mishkovsky M, Frydman L (2008) Progress in hyperpolarized ultrafast 2D NMR spectroscopy. ChemPhysChem 9: 2340-2348
17. Mishkovsky M, Kupče Ē, Frydman L (2007) Ultrafast-based projection-reconstruction three-dimensional nuclear magnetic resonance spectroscopy. J Chem Phys 127: 034507
18. Gal M, Frydman L (2010) Single-scan 2D NMR correlations by multiple coherence transfers. J Magn Reson 203: 311-315
19. Mishkovsky M, Frydman L (2009) Principles and progress in ultrafast multidimensional nuclear magnetic resonance. Annu Rev Phys Chem 60: 429-448
20. Tal A, Frydman L (2010) Single-scan multidimensional magnetic resonance. Prog Nucl Magn Reson Spectrosc 57: 241-292
21. Led JJ, Gesmar H (2010) Fourier Transform and Linear Prediction Methods In: Morris GA, Emsley JW (eds) Multidimensional NMR Methods for the solution state. John Wiley & Sons Ltd.

62

22. Armstrong GS, Cano KE, Mandelshtam VA, Shaka AJ, Bendiak B (2004) Rapid 3D NMR using the filter diagonalization method: application to oligosaccharides derivatized with 13C-labeled acetyl groups. J Magn Reson 170: 156-163
23. 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
24. Meng X, Nguyen BD, Ridge C, Shaka AJ (2009) Enhanced spectral resolution by high-dimensional NMR using the filter diagonalization method and "hidden" dimensions. J Magn Reson 196: 12-22
25. Brüschweiler R (2004) Theory of covariance nuclear magnetic resonance spectroscopy. J Chem Phys 121: 409-414
26. Brüschweiler R, Zhang FL (2004) Covariance nuclear magnetic resonance spectroscopy. J Chem Phys 120: 5253-5260
27. Snyder DA, Bruschweiler R (2009) Generalized indirect covariance NMR formalism for establishment of multidimensional spin correlations. J Phys Chem A 113: 12898-12903
28. Snyder DA, Xu Y, Yang D, Bruschweiler R (2007) Resolution-enhanced 4D 15N/13C NOESY protein NMR spectroscopy by application of the covariance transform. J Am Chem Soc 129: 14126-14127
29. Trbovic N, Smirnov S, Zhang F, Bruschweiler R (2004) Covariance NMR spectroscopy by singular value decomposition. J Magn Reson 171: 277-283
30. Zhang FL, Bruschweiler R (2004) Indirect covariance NMR spectroscopy. J Am Chem Soc 126: 13180-13181
31. Robin M, Delsuc M-A, Guittet E, Lallemand J-Y (1991) Optimized Acquisition and Processing Schemes in Three-Dimensional NMR Spectroscopy. J Magn Reson 92: 645-650
32. Jeannerat D (2007) Computer optimized spectral aliasing in the indirect dimension of H-1-C-13 heteronuclear 2D NMR experiments. A new algorithm and examples of applications to small molecules. J Magn Reson 186: 112-122
33. Coggins BE, Venters RA, Zhou P (2010) Radial sampling for fast NMR: Concepts and practices over three decades. Prog Nucl Magn Reson Spectrosc 57: 381-419
34. Kazimierczuk K, Stanek J, Zawadzka-Kazimierczuk A, Koźmiński W (2010) Random sampling in multidimensional NMR spectroscopy. Prog Nucl Magn Reson Spectrosc 57: 420-434
35. 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
36. Ding KY, Gronenborn AM (2002) Novel 2D triple-resonance NMR experiments for sequential resonance assignments of proteins. J Magn Reson 156: 262-268
37. Koźmiński W, Zhukov I (2003) Multiple quadrature detection in reduced dimensionality experiments. J Biomol NMR 26: 157-166
38. Hiller S, Fiorito F, Wüthrich K, Wider G (2005) Automated projection spectroscopy (APSY). Proc Nat Acad Sci U S A 102: 10876-10881
39. Malmodin D, Billeter M (2005) Signal identification in NMR spectra with coupled evolution periods. J Magn Reson 176: 47-53
40. Kupče Ē, Freeman R (2005) Fast multidimensional NMR: radial sampling of evolution space. J Magn Reson 173: 317-321
41. Kupče Ē, Freeman R (2003) Projection-reconstruction of three-dimensional NMR spectra. J Am Chem Soc 125: 13958-13959
42. Kupče Ē, Freeman R (2004) The Radon transform: A new scheme for fast multidimensional NMR. Concept Magn Reson A 22A: 4-11
43. Kupče Ē, Freeman R (2004) Fast reconstruction of four-dimensional NMR spectra from plane projections. J Biomol NMR 28: 391-395
44. Kazimierczuk K, Koźmiński W, Zhukov I (2006) Two-dimensional Fourier transform of arbitrarily sampled NMR data sets. J Magn Reson 179: 323-328

45. Kazimierczuk K, Zawadzka A, Koźmiński W (2009) Narrow peaks and high dimensionalities: exploiting the advantages of random sampling. J Magn Reson 197: 219-228
46. 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
47. Laue ED, Skilling J, Staunton J, Sibisi S, Brereton RG (1985) Maximum Entropy Method in Nuclear Magnetic Resonance Spectroscopy. J Magn Reson 62: 437-452
48. Hoch JC, Stern AS (1996) NMR Data Processing. Wiley-Interscience, New York
49. Mobli M, Hoch JC (2008) Maximum Entropy Spectral Reconstruction of Nonuniformly Sampled Data. Concepts Magn Reson A 32A: 436-448
50. Luan T, Jaravine V, Yee A, Arrowsmith CH, Orekhov VY (2005) Optimization of resolution and sensitivity of 4D NOESY using multi-dimensional decomposition. J Biomol NMR 33: 1-14
51. Jaravine V, Ibraghimov I, Orekhov VY (2006) Removal of a time barrier for high-resolution multidimensional NMR spectroscopy. Nat Methods 3: 605-607
52. Jaravine VA, Zhuravleva AV, Permi P, Ibraghimov I, Orekhov VY (2008) Hyperdimensional NMR spectroscopy with nonlinear sampling. J Am Chem Soc 130: 3927-3936
53. Szantay C (2008) NMR and the uncertainty principle: How to and how not to interpret homogeneous line broadening and pulse nonselectivity. II. The Fourier connection. Concept Magn Reson A 32A: 1-33
54. Nyquist H (2002) Certain topics in telegraph transmission theory (Reprinted from Transactions of the A. I. E. E., February, pg 617-644, 1928). Proc IEEE 90: 280-305
55. Marion D (2006) Processing of ND NMR spectra sampled in polar coordinates: a simple Fourier transform instead of a reconstruction. J Biomol NMR 36: 45-54
56. Coggins BE, Zhou P (2006) Polar Fourier transforms of radially sampled NMR data. J Magn Reson 182: 84-95
57. Szyperski T, Yeh DC, Sukumaran DK, Moseley HN, Montelione GT (2002) Reduced-dimensionality NMR spectroscopy for high-throughput protein resonance assignment. Proc Natl Acad Sci U S A 99: 8009-8014
58. Malmodin D, Billeter M (2006) Robust and versatile interpretation of spectra with coupled evolution periods using multi-way decomposition. Magn Reson Chem 44: 185-195
59. Coggins BE, Zhou P (2007) Sampling of the NMR time domain along concentric rings. J Magn Reson 184: 207-221
60. 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
61. Kazimierczuk K, Zawadzka A, Koźmiński W (2008) Optimization of random time domain sampling in multidimensional NMR. J Magn Reson 192: 123-130
62. Hyberts SG, Takeuchi K, Wagner G (2010) Poisson-Gap Sampling and Forward Maximum Entropy Reconstruction for Enhancing the Resolution and Sensitivity of Protein NMR Data. J Am Chem Soc 132: 2145-2147
63. Tarczynski A, Allay N, Digital alias-free spectrum estimation of randomly sampled signals, 7th World Multiconference on Systemics, Cybernetics and Informatics, Vol IV, Proceedings, 2003
64. Pannetier N, Houben K, Blanchard L, Marion D (2007) Optimized 3D-NMR sampling for resonance assignment of partially unfolded proteins. J Magn Reson 186: 142-149
65. Press WH, Teukolsky SA, Vetterling WT, Flannery BP (2007) Numerical recipes: the art of scientific computing. Cambridge University Press, Cambridge, UK; New York
66. Davis PJ, Rabinowitz P (1984) Methods of numerical integration. Academic Press, Inc., New York
67. Barthold.E, Ernst RR (1973) Fourier Spectroscopy and Causality Principle. J Magn Reson 11: 9-19
68. Hoch JC, Stern AS, Wagner G (1995) Maximum-Entropy Reconstruction of Nonlinearly Sampled-Data. J Cell Biochem: 76-76

69. 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

70. 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

71. Stern AS, Donoho DL, Hoch JC (2007) NMR data processing using iterative thresholding and minimum l(1)-norm reconstruction. J Magn Reson 188: 295-300

72. Donoho DL (2006) For most large underdetermined systems of linear equations the minimal l(1)-norm solution is also the sparsest solution. Commun Pur Appl Math 59: 797-829

73. Lustig M, Donoho D, Pauly JM (2007) Sparse MRI: The application of compressed sensing for rapid MR imaging. Magn Reson Med 58: 1182-1195

74. Lustig M, Donoho DL, Santos JM, Pauly JM (2008) Compressed sensing MRI. IEEE Signal Process Mag 25: 72-82

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

76. Jiang B, Jiang X, Xiao N, Zhang X, Jiang L, Mao XA, Liu M (2010) Gridding and fast Fourier transformation on non-uniformly sparse sampled multidimensional NMR data. J Magn Reson 204: 165-168

77. Kazimierczuk K, Zawadzka A, Koźmiński W, Zhukov I (2008) 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

78. Kazimierczuk K, Zawadzka-Kazimierczuk A, Koźmiński W (2010) Non-uniform frequency domain for optimal exploitation of non-uniform sampling. J Magn Reson 205: 286-292

79. 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

80. Coggins BE, Venters RA, Zhou P (2004) Generalized reconstruction of n-D NMR spectra from multiple projections: application to the 5-D HACACONH spectrum of protein G B1 domain. J Am Chem Soc 126: 1000-1001

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

82. Werner-Allen JW, Coggins BE, Zhou P (2010) Fast acquisition of high resolution 4-D amide-amide NOESY with diagonal suppression, sparse sampling and FFT-CLEAN. J Magn Reson 204: 173-178

83. Stanek J, Koźmiński W (2010) Iterative algorithm of discrete Fourier transform for processing randomly sampled NMR data sets. J Biomol NMR 47: 65-77

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

85. Keeler J (1984) Elimination of Truncation Artifacts from Nmr-Spectra - Application to C-13 Multiplicity Determination by Two-Dimensional Spectroscopy. J Magn Reson 56: 463-470

86. Shaka AJ, Keeler J, Freeman R (1984) Separation of Chemical-Shifts and Spin Coupling in Proton Nmr - Elimination of Dispersion Signals from Two-Dimensional Spectra. J Magn Reson 56: 294-313

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

88. Barna JCJ, Laue ED, Mayger MR, Skilling J, Worrall SJP (1987) Exponential Sampling, an Alternative Method for Sampling in Two-Dimensional Nmr Experiments. J Magn Reson 73: 69-77

89. Davies SJ, Bauer C, Hore PJ, Freeman R (1988) Resolution Enhancement by Nonlinear Data-Processing - Hogwash and the Maximum-Entropy Method. J Magn Reson 76: 476-493

90. Motáčková V, Nováček J, Zawadzka-Kazimierczuk A, Kazimierczuk K, Žídek L, Šanderová H, Krásný L, Koźmiński W, Sklenář V (2010) Strategy for complete NMR assignment of disordered proteins with highly repetitive sequences based on resolution-enhanced 5D experiments. J Biomol NMR: in press, doi:10.1007/s10858-10010-19447-10853

91. Eberstadt M, Gemmecker G, Mierke DF, Kessler H (1995) Scalar Coupling-Constants - Their Analysis and Their Application for the Elucidation of Structures. Angew Chem Int Ed 34: 1671-1695

92. Griesinger C, Sorensen OW, Ernst RR (1985) Two-Dimensional Correlation of Connected NMR Transitions. J Am Chem Soc 107: 6394-6396

93. Misiak M, Koźmiński W (2007) Three-dimensional NMR spectroscopy of organic molecules by random sampling of evolution time space and multidimensional Fourier transformation. Magn Reson Chem 45: 171-174

94. Misiak M, Koźmiński W, Kwasiborska M, Wójcik J, Ciepichal E, Swiezewska E (2009) Complete H-1 and C-13 signal assignment of prenol-10 with 3D NMR spectroscopy. Magn Reson Chem 47: 825-829

95. Meier S, Benie AJ, Duus JO, Sorensen OW (2009) 3D H2BC: A novel experiment for small-molecule and biomolecular NMR at natural isotopic abundance. J Magn Reson 200: 340-343

96. Meier S, Petersen BO, Duus JO, Sorensen OW (2009) Recent progress in heteronuclear long-range NMR of complex carbohydrates: 3D H2BC and clean HMBC. Carbohyd Res 344: 2274-2278

97. Findeisen M, Bermel W, Berger S (2006) Can one gain spectral information plus time by 3D HMBC via back-projection? Magn Reson Chem 44: 455-458

98. Misiak M, Koźmiński W (2009) Determination of heteronuclear coupling constants from 3D HSQC-TOCSY experiment with optimized random sampling of evolution time space. Magn Reson Chem 47: 205-209

99. Ludwig C, Marin-Montesinos I, Saunders MG, Gunther UL (2010) Optimizing the Polarization Matrix for ex Situ Dynamic Nuclear Polarization. J Am Chem Soc 132: 2508-2509