Available online at www.sciencedirect.com
ScienceDirect
Journal of Magnetic Resonance 192 (2008) 123–130
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Optimization of random time domain sampling in multidimensional NMR

Krzysztof Kazimierczuk, Anna Zawadzka, Wiktor Koźmiński *

Department of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warszawa, Poland

Received 21 December 2007; revised 1 February 2008
Available online 12 February 2008

Abstract

The detailed description of rules for generation of different random sampling schemes is shown and discussed with regard to Multidimensional Fourier Transform (MFT). The influence of different constrained random sampling schedules on FT of constant signal, i.e., Point Spread Function (PSF), is analyzed considering artifacts level and distribution. We found that Poisson disk sampling schedule, which provides a large low-artifact area in the signal vicinity, is the method of choice in the case of nonlinear sampling of time domain in NMR experiments. We have verified the new sampling schemes by application to the 3D HNCACB and ^{15}N-edited NOESY-HSQC spectra acquired for ^{13}C,^{15}N labeled ubiquitin sample.

Keywords: Multidimensional NMR spectroscopy; Multidimensional Fourier Transform; Random data sampling; Proteins; Fast multidimensional data acquisition; Poisson disk; NOESY

1. Introduction

Multidimensional NMR allows for structural studies of biomolecules with atomic resolution. However, in the case of complex molecules spectra are extremely crowded due to the large number of resonances. Therefore, in order to separate signals and enable resonance assignments, one usually adds spectral dimensions. Each of them has its own sampling requirements which determine the number of points needed to achieve desired resolution and spectral width. This leads, independently of overall sensitivity, to very long experiment times.

The goal of new methods of multidimensional NMR is to shorten the experiment time by single scan acquisition [1] or non-conventional, sparse sampling of the evolution time space [2–8], which may be combined with non-FT data processing [9–12].

Generally speaking, lack of information about time domain signal leads to artifacts or peak amplitude/frequency disturbances in spectral domain. There are many methods for processing of sparsely sampled signal. Some of them, like simple additive back projection or Multidimensional Fourier Transform (MFT), admit artifacts, while others give artifact-free spectra with sampling sparseness possibly causing peak frequency or amplitude uncertainty. Sophisticated suppression of artifacts was also employed in some variants of Projection Reconstruction. Various deterministic and non-deterministic approaches to processing of radially sampled data sets and reduction of artifacts were widely discussed before [13].

Spectra obtained by Fourier Transform of sparsely sampled signals compromise signal aliasing and artifacts due to nonlinear sampling. The sampling schedule determines artifact appearance, i.e., Point Spread Function (PSF), like ridges in the case of radial sampling [7,3,5], rings in the case of spiral [7], or concentric ring sampling [4]. According to its definition, shape of PSF does not depend on signal frequencies, and its amplitude is proportional to signal intensity. As we have shown before [8], the relative signal-to-artifact ratio increases proportionally to square root of the number of time domain samples.

* Corresponding author. Fax: +48 22 822 59 96.
E-mail address: kozmin@chem.uw.edu.pl (W. Koźmiński).

1090-7807/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2008.02.003

Recently, we have shown that off-grid random sampling is the best choice because of the lowest artifact level—artifacts appear in each spectral region, but they are very well spread over the frequency domain [8,14]. We have also revealed that even slight modifications of random sampling could substantially improve results [14]. Moreover, such improvement of sampling schedule gives significantly better PSF than obtainable by attempts to increase accuracy of time domain surface integration, which is not useful below Nyquist density [8,14,15]. We have shown that the knowledge of sampling schedule enables for prediction and clean removing of artifacts [14].

In this work we discuss in details the principles of generation of random sampling schedules which predetermine the shape and position of sampling artifacts.

1.1. Random sampling with constraints

Random sampling schedule can be optimized by adding some restrictions on its generation. In our previous paper [14] we presented the idea of jittered stratified sampling. Its algorithm is quite simple. Prior to the generation of points coordinates evolution time space is divided into cells. Then in each cell one point is placed at random position. Such procedure reduces discrepancy of points comparing to unrestricted random sampling schedule.

We confirmed the fact, reported before in works on image processing [16], that using evenly distributed (although still random) data points leads to "blue noise" artifacts of higher frequencies with relatively clean region in the vicinity of the spectral peak. We refer to this region as the "clean area", although it is not completely free-of-artifacts.

Even more optimal sampling can be realized by adding explicit restriction for the minimum relative distances between points. The spectral PSF depends on the "dimensionality" of this restriction and consequently on the shape of histogram showing relative distances between points. Distance constraint R can be defined for any number of dimensions together or separately by condition:

$$\sqrt{\sum_{i=1}^{K}(a_i\Delta t_i)^2} \geqslant R \tag{1}$$

where summation takes place over K dimensions coupled by the distance restriction. Δt_i is the difference between coordinates values in dimension i for a given pair of points. Weighting factors a_i may be added to modify restrictions, which is useful if spectral widths and/or maximum evolution times differ significantly.

Above condition should be satisfied for each pair of points in the sampling schedule. In practice, even most effective algorithms give small amount of points below this limit (if the desired number of samples is required, see below), but it does not affect the general conclusions.

During generation of the sampling schedule any number of such restrictions can be declared. For example, when generating 2D evolution time sampling pattern (i.e., for 3D NMR experiment) one can define the restriction for the dimensions separately:

$$\begin{cases} |\Delta t_1| \geqslant R_1 \\ |\Delta t_2| \geqslant R_2 \end{cases} \tag{2}$$

Or together:

$$\sqrt{(a_1\Delta t_1)^2 + (a_2\Delta t_2)^2} \geqslant R_{12} \tag{3}$$

The comparison of random and random restricted (according to Eq. (3)) sampling is presented in Fig. 1.

There are many algorithms for generation of evolution time values with such point-free sphere around each point

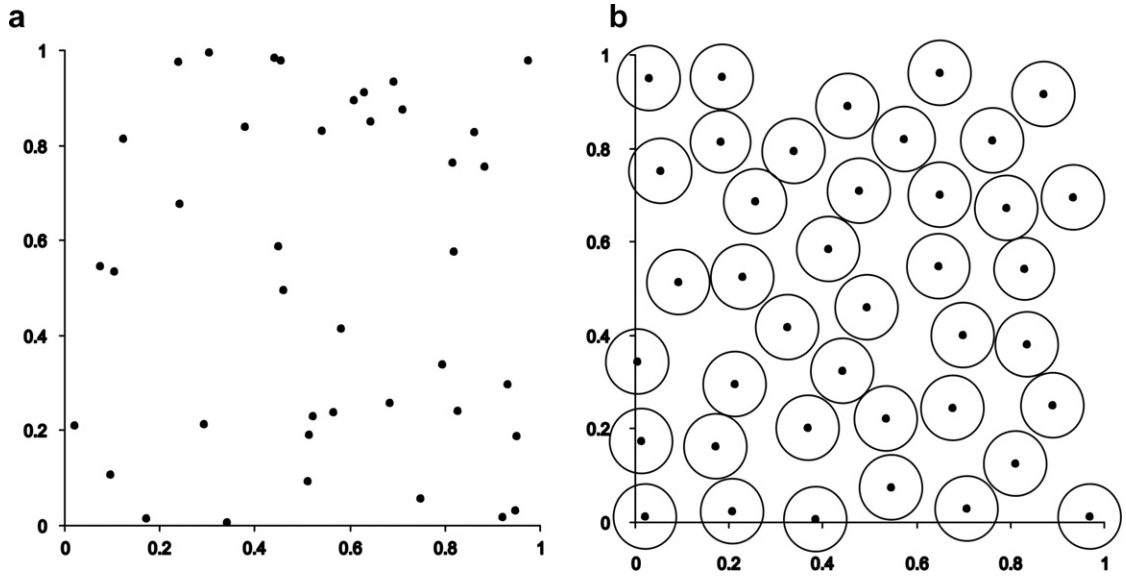

Fig. 1. The illustration of idea of distance-restricted random sampling (point coordinates in arbitrary units): (a) purely random sampling—large discrepancy of points, (b) Poisson disk distribution—even sampling schedule with point-free sphere of diameter R around each sample defined by Eq. (3) with $a_1 = a_2$.

(called Poisson disk sampling), known from computer graphics papers. Some of them are: dart throwing, relaxation dart throwing, Poisson disk tiles, Lloyd's relaxation, tiled blue noise samples, fast hierarchical importance sampling, edge-based Poisson disk tiles, template Poisson disk tiles, corner-based Poisson disk tiles and recursive Wang tiles described in popular reviews [17,18].

They differ mostly in the numerical efficiency and compromise speed and quality of the point response function. For NMR purposes optimization of the generation speed is not the main field of interest as the sampling scheme is generated off-line, before experiment. Thus, the most effective method (in the terms of fulfillment of the distance condition) should be chosen. If one wants to obtain particular number of evolution time points (i.e., defines how long the experiment lasts), which is most common situation, then relaxation dart throwing algorithm is the best choice [19].

This algorithm was discussed in many papers, thus only the most common example of its employment will be presented here. For the generation of the 2D evolution time domain sampling with 2D restriction R_{12} relaxation dart throwing is realized in the following way (see flow-chart presented in Fig. 2):

1. *Number of points* should chosen. Maximum possible distance between points is defined by hexagonal lattice and, for unitary surface $[0,1]\times[0,1]$, could be calculated using expression [17]:

$$R_{\max}=2\sqrt{\frac{1}{2\sqrt{3}N}},\quad \text{where } N \text{ is the number of points.}$$

2. The percentage α of $R_{\max}$ used during generation of sampling schedule should be set. The best results (not too regular and not too high discrepancy) are achieved for α in the range of 65–90%.
3. Generation of points is performed in the loop. Each time the point coordinates are chosen randomly from $[0,1]$ band and the distance restrictions to all points generated before are tested.
4. If the result is successful then the point is accepted.
5. If not, then point coordinates are generated again and again. The number of repetitions (called *relaxation condition*) is declared.
6. If, after declared number of repetitions, point coordinates are still not proper then the distance restriction R is multiplied by a *relaxation multiplier* (which should be chosen from band $(0,1]$ but rather close to 1 like 0.999). The points are generated again in the desired number of repetitions. This procedure is repeated until the point is accepted.
7. After given *number of points* is generated, their coordinates are multiplied by maximum evolution time values.
8. As we shown previously [7,14], it is also convenient to transform such random sampling scheme of uniform points density to decaying points density according to gaussian or exponential distribution. Thus, the initial

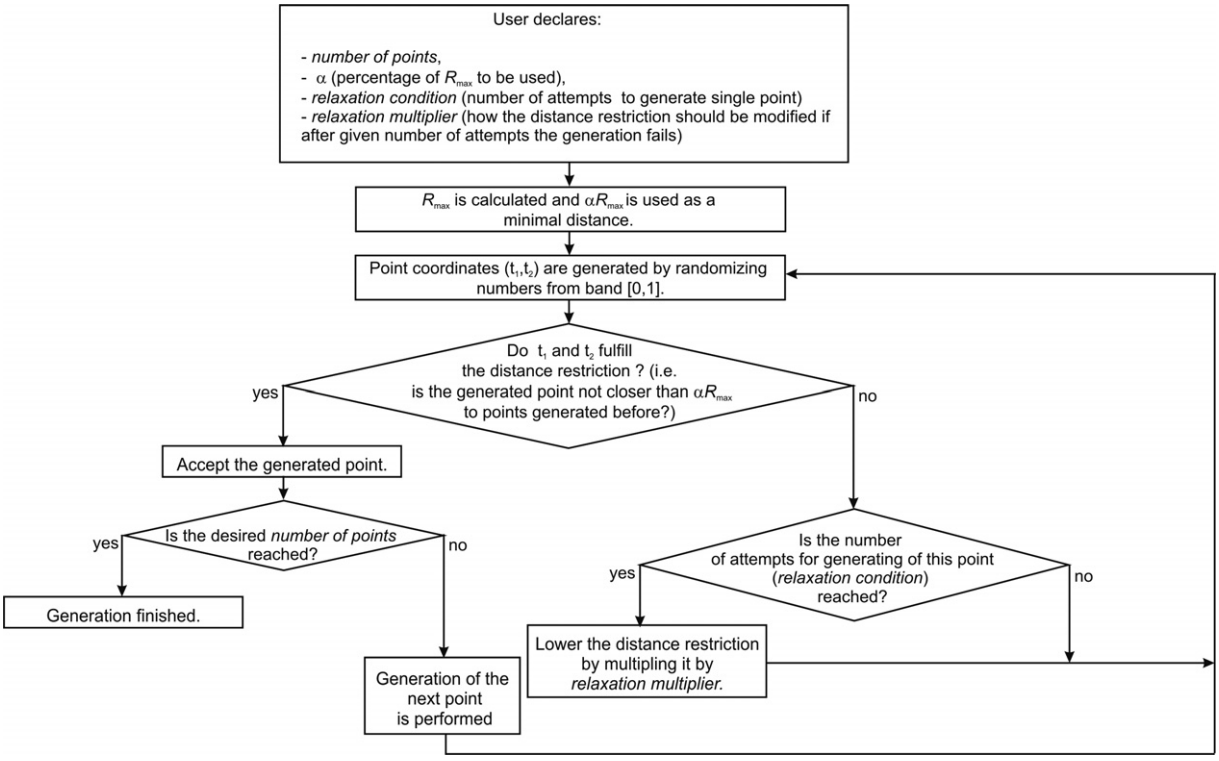

Fig. 2. Flow-chart illustrating procedure of generation of distance-restricted sampling scheme (example of 2D sampling pattern with one 2D distance restriction).

part of interferogram, with high signal-to-noise ratio, is preferred. This is realized by multiplication of point coordinates by appropriate function (still keeping the same evolution time space).

Points are distributed much more evenly than in the strictly random case, and also more evenly than in the case of optimized (or "stratified jittered") sampling discussed in our previous work [14]. Comparison of histograms and integrated PSF for different sampling schedules are plotted in Fig. 3. In the case of Poisson disk sampling (Fig. 3g) most of the points are at about the distance of αR to neighboring points.

1.2. Sampling constraints and spectral point spread function

In all presented cases of random sampling, the distribution of spectral artifacts is strictly connected with the shape of distance restrictions (Fig. 3). If there are no restrictions (i.e. the distribution is purely random) then the artifact level is more or less uniform in each spectral region (Fig. 3b). In other case, there is a clean area in the vicinity of the peak, but then the artifact level raises and reaches maximum for frequencies fulfilling:

$$\sqrt{\sum_{i=1}^{K}\left(\frac{\Delta v_i}{a_i}\right)^2} = \frac{1}{R_{\mathrm{d}}} \tag{4}$$

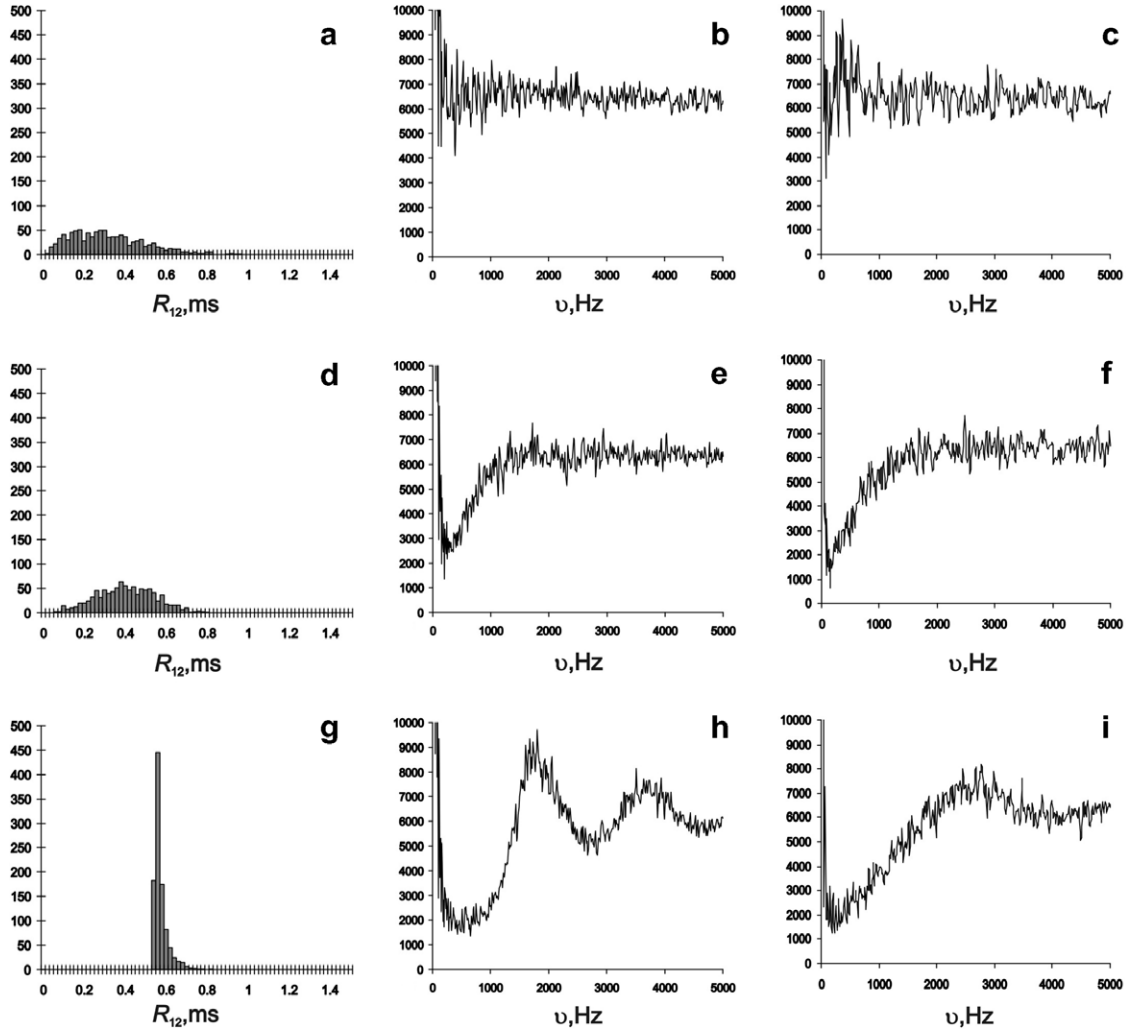

Fig. 3. Simulations presenting dependence of point spread function on the sampling scheme for different random sampling types. 1024 time points and 20 ms × 20 ms evolution time surface were used. Histograms (a), (d), and (g) show number of points with given minimal distance to other points. Pictures (b), (c), (e), (f), (h), and (i) show radial dependence of spectral artifacts and were obtained by integration of modulus of artifact level around the rings in frequency domain. This is justified because of symmetry of artifact level. Plots (b), (e), and (h) are obtained from uniform sampling schedule, while in pictures (c), (f), and (i) points coordinates were transformed to obtain Gaussian distribution of time domain points. In the case of random sampling without distance restriction, i.e., (a), (b), and (c), level of artifacts is more or less equal in the whole spectral domain. Distance histogram of stratified jittered sampling (d) shows maximum at 0.4 ms but it is very broad and thus the artifact-free region in spectrum (e) is very narrow. Poisson disk sampling gives points at more or less the same relative distance about 0.55 ms (g). This gives artifact-free region for small frequencies up to the rapid raise at the 1820 Hz (h), where artifact level reaches maximum and is even higher than in cases (b) and (e). This maximum, however, can be smoothened if decaying points density (e.g., Gaussian distribution) of time coordinates is used (i).

where K is the number of dimensions, the same as in the Eq. (1) and Δv_i is the difference of spectral point coordinate i and a peak coordinate i, and R_d is the dominant distance between points. The level of this maximum could be reduced by application of time decaying points density. This is especially remarkable in the case of Poisson disk sampling (Fig. 3g–i).

It is noteworthy, that according to Eq. (4) minimal distance restriction results in artifacts whose level is appropriate function of distance from the peak in these dimensions where the restriction was declared. Two examples are presented in the Fig. 4c and d. For the given number of points two 1D (Fig. 4c) or one 2D (Fig. 4d) distance restriction can be declared.

In the first case the clean area can be observed only for the one-dimensional spectral region along v_1 and v_2 axis, however, it reaches much greater v_1 and v_2 values. In the second case 2D restriction results in 2D free-of-artifact area around the peak. However, it is not possible to increase it without penalty. The attempt to obtain too restricted Poisson disk sampling converges to regular distribution which causes lines of artifacts parallel to v_1 or v_2 axes. Their level raises with the restricted distance and thus with the clean area. For example, both ridges in Fig. 4d are at the level of 1.5% of the peak height with sampling density equal to 36% of Nyquist density. Employing the elliptic restriction ($a_1 \neq a_2$) differentiates the level of these ridges. Namely, for $a_1 > a_2$ (i.e., broader clean area for v_2) the ridge along v_1 axis is of greater intensity. This case is presented in the experimental examples below, where spectral width significantly differed between two indirectly measured dimensions.

Experimental comparison of HNCACB and ^{15}N-edited NOESY spectra recorded using different random sampling schedules for ^{13}C,^{15}N-labeled ubiquitin is provided in Figs. 5 and 6, respectively. In both cases owing to larger spectral width in F_1 dimension Poisson disk sampling was modified to ellipses ($a_1 > a_2$). Conventional spectrum shown in Fig. 5a suffers from poor resolution, however, when random sampling schedule was used longer maximum

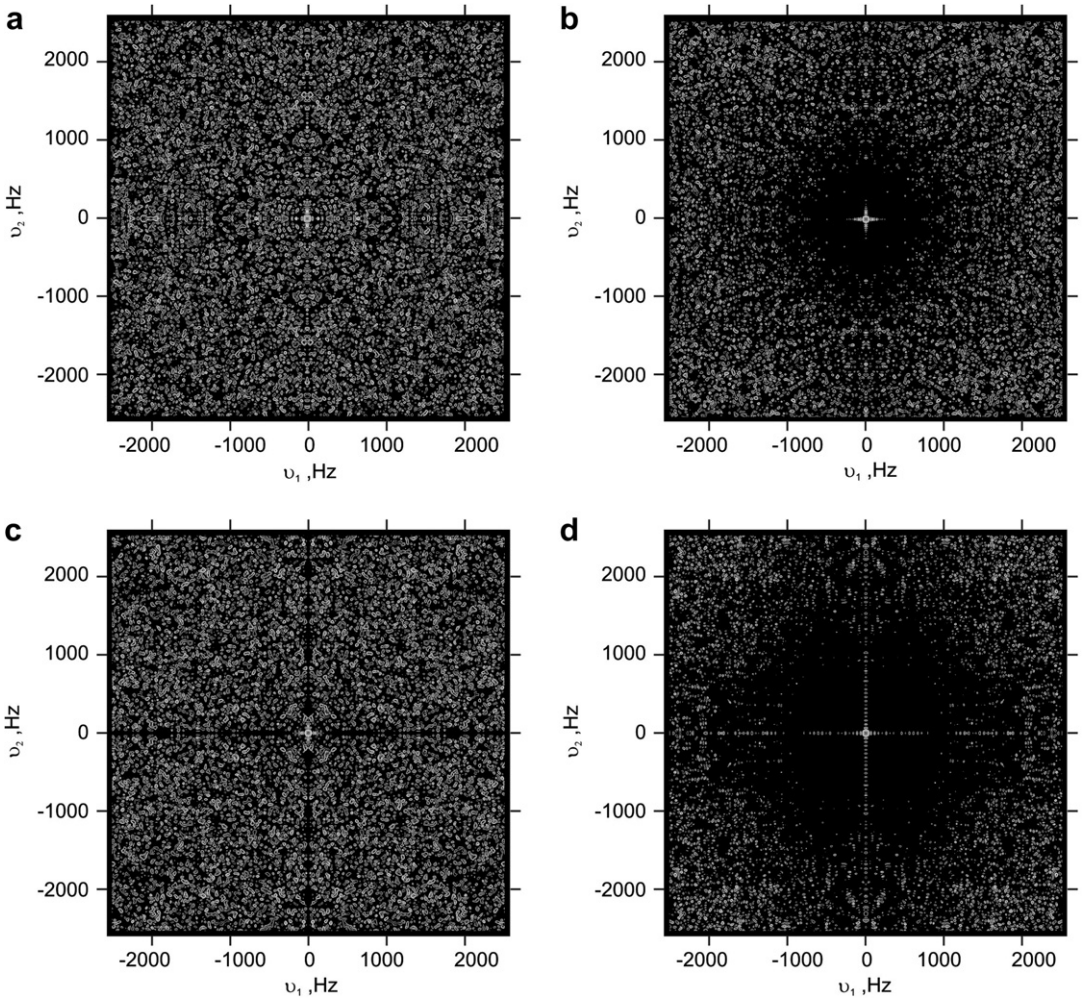

Fig. 4. Point spread functions for different types of random sampling of two-dimensional evolution time space. Positive values are colored gray while negative are white. (a) Random sampling without distance restrictions, (b) stratified jittered sampling, (c) distance-restricted sampling with two separate distance restrictions for two dimensions, and (d) distance-restricted sampling with 2D distance restriction (classical Poisson disk sampling). The same intensity scale was used for all spectra.

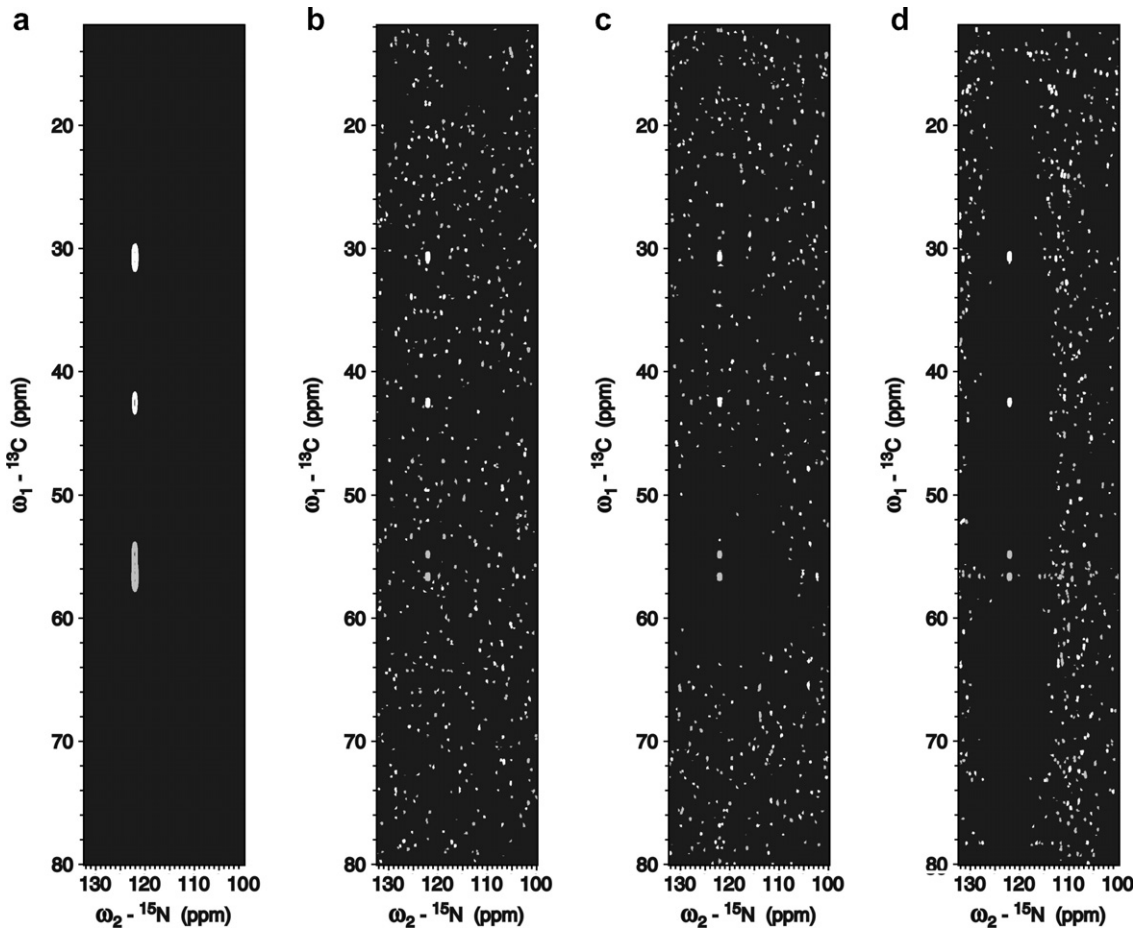

Fig. 5. Contour plots of F_1–F_2 cross sections of 3D HNCACB spectra of human ubiquitin ($\omega_3(^1H) = 8.45$ ppm) for: conventional sampling (a), random sampling without restrictions (b), stratified jittered sampling (c), and Poisson disk sampling (d). Equal intensity threshold, adjusted to the level of artifacts, was used for all plots. The spectral width of 12,000 × 2300 × 12,000 Hz was set in F_1, F_2, and F_3, respectively. 48 × 48 t_1/t_2 data points was recorded in conventional experiment, i.e., the maximum evolution times t_1 and t_2 of 4.0 and 20.9 ms, respectively, were achieved. In the case of experiments with random sampling the maximum evolution times t_1 and t_2 of 25 and 30 ms, respectively, were set (which is equivalent to 11.1% of Nyquist density). Four scans were coherently added in all four data sets for 2304 t_1/t_2 data points, so the acquisition time of both, conventional and randomly sampled experiments, were equal. The spectra were transformed with the resolution of 2048 × 512 × 2048 points in F_1, F_2, and F_3, respectively.

evolution times and thus improved resolution were achieved at the expense of artifacts. While the spectrum plotted in Fig. 5b obtained by unrestricted random sampling reveals uniform artifact distribution, jittered and Poisson disk sampling used for spectra shown in Fig. 5c and 5d provide the clean area in the signals vicinity. Poisson disk samplings give rise to larger clean area and additionally enables optimization of shape of artifact-free regions with respect to specific requirement of the method. Extending the clean area along the ω_2 axis (Fig. 4d) caused higher level of ridge along ω_1. It is, however, still below the level of artifact plateau and thus is not visible. Acquisition of NOESY spectra is more demanding due to high dynamic range of signal intensities. Therefore, using unrestricted random sampling schedule produces uniform PSF which might mask small off-diagonal peaks as shown in Fig. 6a, on the other hand jittered sampling (Fig. 6b), and, especially, optimized Poisson disk sampling (Fig. 6c) schedules enable one to obtain clean area. In the later case the artifact-free region could be adjusted to spectral region of interest, i.e., uncovering small off-diagonal signals. Very wide clean area along ω_1 (Fig. 6c) was achieved at the expense of higher ridge in the ω_2 dimension. These residual artifacts could be easily removed employing cleaning procedure described in Ref. [14]. The result showing clean spectrum is drawn in Fig. 6d.

2. Experimental

The 3D HNCACB and ^{15}N-edited NOESY-HSQC spectra were recorded for 1.5 mM ^{13}C, ^{15}N-double labeled human ubiquitin in 9:1 H_2O/D_2O at pH 4.5 at 298 K on a Varian NMR System 700 spectrometer equipped with a Performa XYZ PFG unit and using the 5 mm ^{1}H, ^{13}C, ^{15}N-triple resonance probehead with high power ^{1}H, ^{13}C, and ^{15}N $\pi/2$ pulses of 5.9, 13.5, and 31.0 μs, respectively. The pulse sequences were adapted from the Varian Userlib BioPack package. The MFT was performed employing PC with 3.0 GHz Pentium 4 processor running under Linux OS. The software is available from authors.

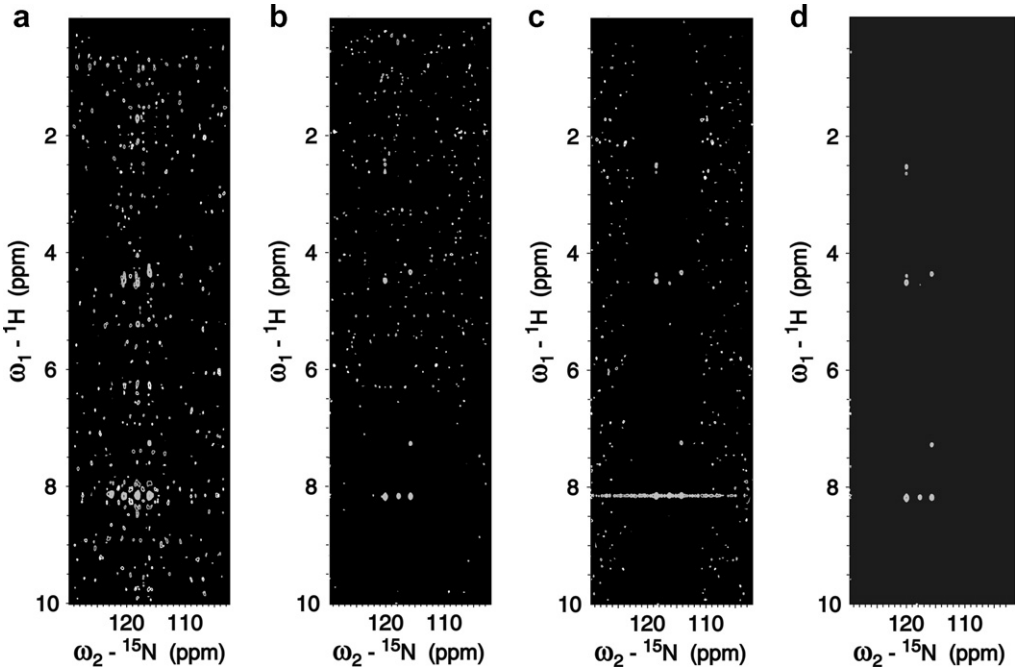

Fig. 6. Contour plots of F_1–F_2 cross sections of 3D NOESY-HSQC spectra of human ubiquitin ($\omega_3(^1H) = 8.162$ ppm) for: random sampling without restrictions (a), stratified jittered sampling (b), and Poisson disk sampling (c). Additionally, artifact level can be significantly reduced (d) by employing cleaning procedure discussed in our previous work [14]. All spectra are plotted in the same intensity scale. The spectral width of 12,000 × 2300 × 12,000 Hz was set in F_1, F_2, and F_3, respectively. The maximum evolution times t_1 and t_2 were both set to 30 ms (thus the relative density of points was 9.3% of Nyquist density). Four scans were coherently added in all four data sets for 2304 t_1/t_2 data points. The spectra were transformed with the resolution of 2048 × 512 × 2048 points in F_1, F_2, and F_3, respectively.

3. Conclusions

We have shown that it is possible to shift artifacts, rising due to irregular sampling of evolution time space, away from the peaks. This is realized by defining some distance constraints during generation of sampling schedule. The best and general approach to generation of distance-restricted sampling pattern is relaxation dart throwing used for Poisson disk sampling, well known in image processing. It is originally developed for 2D but can be easily extended to higher dimensions. Also distance restriction can be easily modified and separated to restrictions of lower dimensionality.

In our opinion distance-restricted sampling is the best option of random sampling, and the optimal results can be obtained if dimensionality of the restriction is equal to the dimensionality of evolution time space. In some experiments, however, having artifact-free-zone along spectral axes may become more convenient.

In the case of significant differences in spectral width in two dimensions involved in the minimum distance condition it is convenient to modify Poisson disk to "Poisson Ellipse" distribution.

Acknowledgments

This work was supported by Grant Number: N301 07131/2159, founded by Ministry of Science and Higher Education in years 2006–2009. The NMR measurements were accomplished at the Structural Research Laboratory, Chemistry Department, University of Warsaw, Poland. The 700 MHz Varian NMR Systems spectrometer was founded by Grant: 217/FNiTP/115/2005 of Foundation of Polish Science and Technology.

References

[1] L. Frydman, T. Scherf, A. Lupulescu, The acquisition of multidimensional NMR spectra within a single scan, Proc. Natl. Acad. Sci. USA 99 (2002) 15662–15858.
[2] Ē. Kupče, R. Freeman, Projection-reconstruction of three-dimensional NMR spectra, J. Am. Chem. Soc. 125 (2003) 13958–13959.
[3] B.E. Coggins, P. Zhou, Polar Fourier transforms of radially sampled NMR data, J. Magn. Reson. 182 (2006) 84–95.
[4] B.E. Coggins, P. Zhou, Sampling of the NMR time domain along concentric rings, J. Magn. Reson. 184 (2007) 207–221.
[5] D. Marion, Processing of ND NMR spectra sampled in polar coordinates: a simple Fourier transform instead of reconstruction, J. Biomol. NMR 36 (2006) 45–54.
[6] D. Marion, Fast acquisition of NMR spectra using Fourier transform of non-equispaced data, J. Biomol. NMR 32 (2005) 141–150.
[7] K. Kazimierczuk, W. Koźmiński, I. Zhukov, Two-dimensional Fourier transform of arbitrarily sampled NMR data sets, J. Magn. Reson. 179 (2006) 323–328.
[8] K. Kazimierczuk, A. Zawadzka, W. Koźmiński, I. Zhukov, Random sampling of evolution time space and Fourier transform processing, J. Biomol. NMR 36 (2006) 157–168.
[9] V. Orekhov, I. Ibraghimov, M. Billeter, Optimizing resolution in multidimensional NMR by three-way decomposition, J. Biomol. NMR 27 (2003) 165–173.

[10] D. Malmodin, M. Billeter, Signal identification in NMR spectra with coupled evolution periods, J. Magn. Reson. 176 (2005) 47–53.
[11] D. Rovnyak, D.P. Frueh, M. Sastry, Z.-Y.J. Sun, A.S. Stern, J.C. Hoch, G. Wagner, Accelerated acquisition of high resolution triple-resonance spectra using non-uniform sampling and maximum entropy reconstruction, J. Magn. Reson. 170 (2004) 15–21.
[12] D.A. Snyder, F. Zhang, R. Brueschweiler, Covariance NMR in higher dimensions: application to 4D NOESY spectroscopy of proteins, J. Biomol. NMR 39 (2007) 165–175.
[13] J.W. Yoon, S. Godsill, Ē. Kupče, R. Freeman, Deterministic and statistical methods for reconstructing multidimensional NMR spectra, Magn. Reson. Chem. 44 (2006) 197–209.
[14] K. Kazimierczuk, A. Zawadzka, W. Koźmiński, I. Zhukov, Lineshapes and artifacts in Multidimensional Fourier Transform of arbitrary sampled NMR data sets, J. Magn. Reson. 188 (2007) 344–356.
[15] N. Pannetier, K. Houben, L. Blanchard, D. Marion, Optimized 3D NMR sampling for resonance assignment of partially unfolded proteins, J. Magn. Reson. 186 (2007) 142–149.
[16] D.P. Mitchell, Generating antialiased images at low sampling densities, in: Computer Graphics (Proceedings of ACM SIGGRAPH 87) 21, 4, 1987, pp. 65–72.
[17] A. Lagae, P. Dutré, A Comparison of methods for generating Poisson disk distributions, Computer Graphics Forum (OnlineEarly Articles), in press. doi:10.1111/j.1467-8659.2007.01100.x.
[18] R.V. Klassen, Filtered Jitter, Comput. Graph. Forum 19 (2000) 223–230.
[19] M. McCool, E. Fiume, Hierarchical Poisson disk sampling distributions, in: Graphics Interface '92, 1992, pp. 94–105.