Published in final edited form as:

J Magn Reson. 2015 May ; 254: 121–130. doi:10.1016/j.jmr.2015.02.015.

Nonuniform sampling of hypercomplex multidimensional NMR experiments: Dimensionality, quadrature phase and randomization

Adam D Schuyler^{a,*}, Mark W Maciejewski^a, Alan S Stern^b, and Jeffrey C Hoch^{a,*}

^aDepartment of Molecular Biology and Biophysics, UConn Health, 263 Farmington Avenue, Farmington, CT 06030-3305, USA

^bRowland Institute at Harvard, 100 Edwin H. Land Boulevard, Cambridge, MA 02142, USA

Abstract

Nonuniform sampling (NUS) in multidimensional NMR permits the exploration of higher dimensional experiments and longer evolution times than the Nyquist Theorem practically allows for uniformly sampled experiments. However, the spectra of NUS data include sampling-induced artifacts and may be subject to distortions imposed by sparse data reconstruction techniques, issues not encountered with the discrete Fourier transform (DFT) applied to uniformly sampled data. The characterization of these NUS-induced artifacts allows for more informed sample schedule design and improved spectral quality. The DFT–Convolution Theorem, via the point-spread function (PSF) for a given sampling scheme, provides a useful framework for exploring the nature of NUS sampling artifacts. In this work, we analyze the PSFs for a set of specially constructed NUS schemes to quantify the interplay between randomization and dimensionality for reducing artifacts relative to uniformly undersampled controls. In particular, we find a synergistic relationship between the indirect time dimensions and the “quadrature phase dimension” (i.e. the hypercomplex components collected for quadrature detection). The quadrature phase dimension provides additional degrees of freedom that enable partial-component NUS (collecting a subset of quadrature components) to further reduce sampling-induced aliases relative to traditional full-component NUS (collecting all quadrature components). The efficacy of artifact reduction is exponentially related to the dimensionality of the sample space. Our results quantify the utility of partial-component NUS as an additional means for introducing decoherence into sampling schemes and reducing sampling artifacts in high dimensional experiments.

Keywords

aliasing; partial-component NUS; discrete Fourier transform (DFT); point-spread function (PSF); compressed sensing

© 2015 Published by Elsevier Inc.

*Corresponding Author: schuyler@uchc.edu | T: +1 (860) 679-1496 | F: +1 (860) 679-3408. *Corresponding Author: hoch@uchc.edu | T: +1 (860) 679-3566 | F: +1 (860) 679-3408.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Introduction

Nonuniform sampling (NUS) methods, and associated spectral reconstruction techniques, are increasingly used to reduce sampling requirements in multidimensional NMR. In seminal work by Barna and colleagues, data from two-dimensional NMR experiments collected using an exponentially biased selection of t_1 values from a Cartesian grid spaced at the Nyquist interval [1], were processed with maximum entropy reconstruction [2, 3] to compute the spectrum. Off-grid NUS methods employing sampling along radial vectors in the indirect time dimensions were introduced in the 2000s by Ding and Gronenborn [4], Kupce and Freeman [5] and Kim and Szyperski [6]. Spectral estimation methods such as back-projection reconstruction [7, 8] and the G-matrix Fourier transform [6, 9], were introduced to handle off-grid radial sampling. Because they utilize different spectral reconstruction techniques, the connection between the on-grid approach of Barna et al. and the off-grid radial sampling approaches was not immediately recognized. However, Mobli et al. [10] demonstrated the close connection by using maximum entropy reconstruction for radial sampling schemes that fall on a Cartesian grid.

The principle applications of NUS to date have been to obtain high resolution spectra while minimizing experiment time (see Maciejewski et al. [11] and Mobli et al. [12] for review). The omission of points from a uniform sampling grid results in gaps which, according to the Nyquist theorem, introduce aliased peaks that appear as artifacts in the final spectrum. From the earliest work by Barna et al., it has been clear that the distribution of sample times influences the distribution and magnitude of sampling artifacts. Understanding these artifacts is an important first step for improving the quality of spectra obtained from NUS experiments. Critical comparison of different sampling strategies is complicated by the non-linearity of most non-Fourier methods of spectral reconstruction. Nevertheless, a growing body of empirical evidence, coupled with theoretical insights, has yielded two fundamental principles for the design of efficient sampling schemes. The first is that randomness is important for minimizing sampling artifacts [13, 14, 15, 16]. The second is that tailoring the sampling distribution to capture more samples at times when the signal envelope is larger and less samples when it is smaller helps improve sensitivity [1, 17]. Beyond these general principles, more specific prescriptions have remained elusive because the quality of spectra obtained using NUS depends not only on the chosen NUS schedule, but also on the nature of the signals (e.g. noise level, dynamic range and signal decay rates), the dimensionality of the experiment and the method used to reconstruct the spectrum.

The point-spread function (PSF) is the discrete Fourier transform (DFT) of the sampling function, which is a multidimensional array with an element equal to one for each free induction decay (FID) that is sampled and equal to zero for each FID that is not sampled. Schmieder et al. [18] used the PSF as a quantitative tool for comparing sampling schemes, and despite the observation by Lustig et al. [19] that the PSF is a “natural tool to measure incoherence” of sampling schemes, the PSF has only served a minor supporting role in the investigations of NUS [20, 21, 22]. In the present work we utilize the peak-to-sidelobe ratio (PSR, an adapted form of the sidelobe-to-peak ratio of Lustig et al. [19]) which is the ratio between the magnitude of the zero-frequency component and the largest satellite (non-zero-

frequency component) in the PSF. The PSR serves as a quantitative measure of the coherence among the sampled times in the sampling function. As such, PSR is an *a priori* measure akin to “signal-to-noise”, in that it gives the upper bound on the ratio between the zero-frequency component and the largest NUS induced artifact expected for a given sampling function. The artifacts follow the upper bound when the NUS data is zero-augmented (i.e. FIDs not collected by the NUS schedule are zero filled) and processed by DFT, whereas the artifacts are reduced when spectral reconstruction methods make no assumptions about the missing FIDs (e.g. maximum entropy [23, 24]).

We present a PSR analysis of several carefully constructed sampling schemes designed to elucidate the role of dimensionality and randomization in NUS. These schemes introduce randomization along time and/or quadrature phase dimensions. The sampling schemes developed here are not intended for use in NMR experiments, but they do provide a useful perspective on the importance of decoherence and its relation to dimensionality. Our results reveal the utility of quadrature phase as an additional degree of freedom, through which randomization can further reduce coherence in NUS schemes and thereby reduce sampling artifacts relative to schedules which do not sample the quadrature phase.

Theory

Spectra for NUS data collected on a Cartesian grid spaced at the Nyquist interval are typically estimated using non-Fourier methods that suppress sampling artifacts relative to Fourier methods applied to zero-augmented data. The ability of these methods to suppress artifacts is subject to limitations, principally because of experiment noise. However, the DFT of NUS data (where zeros are used to augment the samples missing from the uniform grid) is a convenient tool for characterizing the relative performance of different sampling schemes because of its particularly simple relationship to the spectrum obtained by DFT of uniformly sampled data. The DFT–Convolution Theorem states that the DFT of zero-augmented NUS data is given by the convolution of the PSF with the DFT spectrum of the corresponding uniformly sampled data set.

Quadrature detection typically used to determine the sign of spectral frequencies requires separate, sequential experiments when used along indirect time dimensions (as opposed to the direct acquisition dimension, where in-phase and out-of-phase detection can be performed simultaneously) [25]. The majority of NUS schemes used to date collect all 2^d quadrature components for each sampled time point for quadrature detection conducted along d indirect dimensions. We recently described random phase detection (RPD, [26]), in which only a single quadrature component is randomly selected for detection from among the 2^d quadrature components, enabling a factor of 2^d reduction in the number of FIDs collected per time index, relative to conventional quadrature detection. RPD is one example of *partial-component NUS*, which as a class, includes any scheme that detects less than 2^d quadrature components for a sampled time point. *Full-component NUS* collects all 2^d components for each sampled time.

Partial-component NUS makes the relationship between the DFT of the zero-augmented NUS data and the PSF more complicated. The DFT–Convolution Theorem no longer

applies, and there is no longer a single-valued sampling function from which the PSF can be computed. Instead there are separate sampling functions for each quadrature component of the hypercomplex data. As shown in our previous work [27], the DFT spectrum of zero-augmented partial-component NUS data is given by a linear combination of convolutions, one for each of the sampling functions; a corresponding partial-component PSF may be computed as the aggregate power of the PSFs for the individual sampling functions. In our notation for partial-component NUS, the entries in a d -dimensional partial-component sampling function (for a $(d + 1)$ -dimensional experiment) are defined by

$$S[k_1, \dots, k_d]\{\phi\} \quad (1)$$

where $k_i \in \{1, \dots, m_i\}$ is the index along indirect dimension i up to a maximum increment of m_i and ϕ is the hypercomplex component taken from $\mathcal{P}_d$, which is the set of all 2^d hypercomplex components on d -dimensions. With the real and imaginary components along each dimension referred to as “R” and “I”, respectively, we have, for example, $\mathcal{P}_2 = \{\text{“RR”}, \text{“RI”}, \text{“IR”}, \text{“II”}\}$. The k_i values index along the indirect time dimensions and the ϕ value indexes along the “quadrature phase dimension”. Entries in S with a value of 1 indicate FIDs that are collected and values of 0 indicate FIDs that are not. Sample coverage (c) is the percentage of FIDs from a uniform sampling grid that are collected by a sample function and is computed as

$$c = \frac{\sum_{k_1, \dots, k_d, \phi} S[k_1, \dots, k_d]\{\phi\}}{2^d \cdot (m_1 \dots m_d)} \quad (2)$$

If a desired sample coverage equates to a non-integer number of FIDs, the number of FIDs collected is rounded up by one.

Methods

Three types of NUS schemes are considered here: random sampling (R), constant-offset undersampling ($\tilde{}$) and random-offset undersampling ($\tilde{}$). Each of these types is deployed as either full-component (“F” subscript) or partial-component (“P” subscript). The random sample schedules (R_F and R_P) are generated by the random selection of sample points across the uniform sample grid. The constant-offset ($\tilde{R}_F$ and $\tilde{R}_P$) and random-offset ($\tilde{\tilde{R}}_F$ and $\tilde{\tilde{R}}_P$) undersampling schedules are constructed as hypercomplex $(d - 1)$ -dimensional arrays containing copies of a uniformly un-dersampled 1D vector, each placed with an index offset along the first indirect dimension; these schedules are collectively referred to as 1D-generated (1DG). With respect to the notation established in Equation 1, the index along each 1D vector is given by the time index $[k_1]$ and the location of each 1D vector within the hypercomplex $(d-1)$ -dimensional array is given by the remaining time indices $[k_2, \dots, k_d]$ along with the “quadrature phase index” $\{\phi\}$.

Example schedules in two indirect dimensions are shown in Figure 1 for the full-component schedules (R_F , $\tilde{R}_F$, $\tilde{\tilde{R}}_F$) and Figure 2 for the partial-component schedules (R_P , $\tilde{R}_P$, $\tilde{\tilde{R}}_P$); each with its corresponding PSF power. The sampling schemes are defined as

$$\begin{aligned}
R_F[k_1, \dots, k_d]\{\phi\} &= \sigma^c[k_1, \dots, k_d] \\
R_P[k_1, \dots, k_d]\{\phi\} &= \sigma^c[k_1, \dots, k_d]\{\phi\} \\
\bar{U}_F[k_1, \dots, k_d]\{\phi\} &= \mathbf{u}^p[k_1] \\
\bar{U}_P[k_1, \dots, k_d]\{\phi\} &= \mathbf{u}^p[k_1 - \text{rand}^p\{\phi\}] \\
\tilde{U}_F[k_1, \dots, k_d]\{\phi\} &= \mathbf{u}^p[k_1 - \text{rand}^p[k_2, \dots, k_d]] \\
\tilde{U}_P[k_1, \dots, k_d]\{\phi\} &= \mathbf{u}^p[k_1 - \text{rand}^p[k_2, \dots, k_d]\{\phi\}]
\end{aligned} \tag{3}$$

The random sample schedules (R_F and R_P) draw values from the function σ^c , which uses MATLAB's combined multiple recursive random number generator (mrg32k3a), to return a value of 0 or 1, under the constraints that c percent of all sample points are selected and an equal distribution across all hypercomplex components is maintained (i.e. the difference in the number of sample points taken across any two hypercomplex components is no greater than 1). The σ^c function for R_F takes the time indices $[k_1, \dots, k_d]$ as input parameters, but not the quadrature phase $\{\phi\}$, indicating that the output value depends on the combination of time indices given as input parameters, but is independent of quadrature phase (i.e. full-component). The σ^c function for R_P takes the time indices $[k_1, \dots, k_d]$ and quadrature phase as input parameters $\{\phi\}$, thereby producing a partial-component schedule.

The uniformly undersampled 1D vector used to generate the 1DG schedules in Equation 3 is defined as

$$\mathbf{u}^p[k_1] = \begin{cases} 1 & \text{mod}(k_1 - 1, p) = 0 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where the undersampling period is given by $p \in \{2, 3, \dots\}$ and k_1 indexes the elements along the vector. The mod operator selects one sample point out of every p and the $k_1 - 1$ index shift forces the first point to be sampled in each period rather than the last. The copies of $\mathbf{u}^p$ used to fill a hypercomplex $(d - 1)$ -dimensional array for each of the 1DG schedule types may be placed such that the first sample point taken along *every* vector occurs at $t_1 = 1$ (i.e. constant-offset) or the offset of the first sample point may be randomly selected for *each* vector (i.e. random-offset). This random offset is controlled by the rand^p function, which uses MATLAB's combined multiple recursive generator (mrg32k3a) to return a value from the set of integers $\{0, \dots, p-1\}$. As similarly defined for the σ^c function, the rand^p function may take time indices and/or quadrature phase as input parameters, with the output value only dependent on the specified input parameters and independent of any omitted parameters.

The construction of each 1DG schedule is discussed below, with references made to features observed in the 2D example schedules shown in Figures 1 and 2. The concepts generalize to arbitrary dimensionality.

F sampling (Figure 1B) has a fixed offset along t_1 ; the first sample point along every t_1 column is taken at $t_1 = 1$. The same sampling function is taken across each hypercomplex component (i.e. full-component). The rows along t_2 are either completely sampled or not sampled at all. As evidenced by the row of perfectly aliased peaks along f_1 flanking the zero-

frequency component at $f_1 = f_2 = 0$ in the PSF power, this schedule type corresponds to coherent undersampling.

$\tilde{U}_F$ sampling (Figure 1C) has a random offset along t_1 ; the first sample point along each t_1 column is randomly selected from the set $\{1, \dots, p\}$. The same sampling function is taken across each hypercomplex component (i.e. full-component). The random offset along t_1 allows the uniformly undersampled generating vector to be placed so that the t_1 columns are free to translate up and down relative to each other, leading to decoherence of the sampling across t_2 . As a result, the perfectly aliased peaks observed in U_F are distributed across f_2 , while the zero-frequency component remains.

U_P sampling (Figure 2B) has a fixed offset along t_1 . However, unlike U_F which has the same fixed offset for each 1D vector along t_1 , U_P allows an independent selection of a fixed offset for each quadrature phase (i.e. partial-component). Despite the independent selection of the offset value for each quadrature phase, the rows along t_2 are either completely sampled or not sampled at all resulting in the same perfectly aliased peaks in the PSF power as observed in U_F .

$\tilde{U}_P$ sampling (Figure 2C) has a random offset along t_1 , which is also taken independently across quadrature phase (i.e. partial-component). This allows each t_1 column for each hypercomplex component to be placed with a random offset. All t_1 columns are independent producing decoherence across all t_2 rows. The aliased peaks in the PSF power are attenuated in a similar fashion to $\tilde{U}_F$ sampling.

PSR achieves a low value of 1 for an NUS-induced artifact of the same intensity as the zero-frequency component of the PSF (i.e. perfect aliasing). At the other extreme, PSR approaches infinity as sampling artifacts in the PSF are reduced to zero intensity; a condition reached in the limit as sample coverage approaches 100%. These two extremes are captured in the present study by U_F and R_F , respectively. The analysis that follows explores how the sampling artifacts of U_F are reduced in response to randomization of the undersampling offset: (i) along the indirect time dimensions ($\tilde{U}_F$, Figure 1C), (ii) along the “quadrature phase dimension” (U_P , Figure 2B) and (iii) along both the indirect time and “quadrature phase” dimensions simultaneously ($\tilde{U}_P$, Figure 2C). The success of $\tilde{U}_F$, U_P and $\tilde{U}_P$ are quantified by their PSR performance relative to the extremes established by U_F and R_F sampling. In addition, R_P sampling is included so that the benefits of partial-component sampling may be quantified by comparing full- and partial-component forms of each sample schedule type (i.e. R_F vs R_P , U_F vs U_P and $\tilde{U}_F$ vs $\tilde{U}_P$). The PSR analysis specifically focusses on the role of sample schedule dimensionality.

Results and Discussion

PSR Analysis

The magnitude of sampling artifacts expected for each of the 6 NUS schedules considered here is determined by computing PSR values of sample schedules for a varying number of indirect dimensions (d), maximum increment along each dimension (m) and sample

coverage (c). Note that the simulations are performed with the same maximum increment along each dimension. Therefore, the total number of time points in the sample grid is

$$g=m^d \quad (5)$$

There are four preliminary items to note: (1) The sampling schemes considered in this work select an equal distribution of FIDs among the quadrature components. Consequently, the general expression for sample coverage given in Equation 2, simplifies to a percentage of the g FIDs selected from the sample grid on each quadrature component. Accordingly, the grid size parameter g is commonly employed as the horizontal axis in the PSR analyses that follow. (2) The sample coverage for the 1DG schedules is determined by the periodicity of the undersampling as $c = 1/p$. When comparing 1DG schedules with the random schedules, the inverse relationship between coverage and periodicity is employed to achieve a common horizontal axis. (3) When a periodicity value is not a divisor of the maximum increment, it is impossible to achieve a coverage of exactly $c = 1/p$. In addition, if a random offset is employed (e.g. $\tilde{U}_F$ or $\tilde{U}_P$), the number of FIDs collected in any 1D vector along the uniformly undersampled dimensions may vary by at most one. (4) All schedules, except F , include randomization and are repeated 50 times for the parameter combinations listed in Table 1 and the results averaged to determine mean PSR values. The resulting PSR data is shown for random sampling (R_F and R_P) in Figure 3; constant-offset undersampling (F and P) in Figure 4; and random-offset-undersampling ($\tilde{U}_F$ and $\tilde{U}_P$) in Figure 5. The PSR plots are on log-log axes with the horizontal axis giving the number of time points in the sample grid (g). The standard deviations for each data point in Figures 3–5 are reported in Supplemental Tables S1–S6. The corresponding error bars are not shown in Figures 3–5, as the standard deviations are small enough such that the span of the error bars is narrower than the height of the symbols marking each data point.

Constant-offset undersampling (F and P) exhibits perfect aliases, resulting in $PSR \approx 1$; the F and P data points deviating from $PSR=1$ are addressed in the “Periodicity of the DFT” section. In contrast, the PSR data for random sampling (R_F and R_P) and random-offset undersampling ($\tilde{U}_F$ and $\tilde{U}_P$) are greater than one. In particular, these data form collinear tiers for each combination of dimensionality and sample coverage, with the slope of each tier varying only with dimensionality (i.e. on each plot, all lines of the same color have the same slope). These observations are described by the exponential relationship

$$PSR \sim g^\alpha = m^{\alpha \cdot d} \quad (6)$$

where α is the slope of each line. This exponential relationship shows that the slope (α) of the linear data in the log-log PSR plots acts as a scaling factor on dimensionality (d) in the exponent. In other words, adding a dimension results in the number of time points in the grid increasing by a factor of m , while the PSR value increases by a factor of m^α . These relationships allow for the single scalar α to characterize how random sampling (R_F and R_P) and random-offset undersampling ($\tilde{U}_F$ and $\tilde{U}_P$) utilize dimensionality to increase PSR by reducing artifacts relative to constant-offset undersampling (F and P).

The role of dimensionality is quantitatively characterized in Figure 6 by computing an average α value for each sample schedule type at each value of dimensionality (d). The constant-offset undersampling schedules ($\bar{F}$ and $\bar{P}$) are not randomized and exhibit perfect aliases; these bars are at $\alpha \approx 0$ and are barely visible in the figure. At the other extreme, the random sampling schedules are randomized across all indirect dimensions and achieve $\alpha \approx 1$. The random-offset undersampling schedules utilize randomization in $d - 1$ of the d indirect time dimensions and achieve $\alpha \approx \frac{d-1}{d}$. In essence, the α value is a measure of the randomness or decoherence among the nonuniformly sampled indirect time dimensions.

The α values in Figure 6 reveal an advantage for partial-component schedules relative to traditional full-component NUS, albeit secondary to the role of randomization and dimensionality. The slight increase in α values for partial-component sampling indicates an effective usage of the quadrature phases to introduce sampling decoherence. In addition to this performance boost of the exponential relationship, Figure 7 gives a direct assessment of the merits of partial-component NUS. Figure 7A plots the ratio of PSR data for R_P relative to R_F (i.e. right panel divided by left panel from Figure 3) and Figure 7B plots the ratio of PSR data for $\tilde{U}_P$ relative to $\tilde{U}_F$ (i.e. right panel divided by left panel from Figure 5). Each panel in the resulting Figure 7 shows a linear increase with dimensionality – an additional scaling factor that acts independently of the exponential term.

Modifying Equation 6, to include the substitution $\alpha = 1$ for R_F and R_P ; the substitution $\alpha \approx \frac{d-1}{d}$ for $\tilde{U}_F$ and $\tilde{U}_P$ and the scaling factor of d for R_P and $\tilde{U}_P$, produces the following expressions describing how PSR relates to dimensionality

$$PSR \approx \begin{cases} 1 & \bar{U}_F, \bar{U}_P \\ m^{d-1} & \tilde{U}_F \\ d \cdot m^{d-1} & \tilde{U}_P \\ m^d & R_F \\ d \cdot m^d & R_P \end{cases} \quad (7)$$

Periodicity of the DFT

Close inspection of Figures 4 and 5 reveals two unexpected results: (1) the PSR values are higher for $c = 20\%$ than for $c = 50\%$; and (2) the $m = 40$ at $c = 20\%$ data points are diminished relative to the collinear data points collected at the other dimension lengths. These curious findings are related to the periodicity of the DFT.

A $\bar{F}$ schedule of dimension length m and period length p , has a PSF power with a row of p equally spaced peaks along the undersampled dimension (Figure 8). The middle peak is the zero-frequency component and the remaining peaks are aliases introduced by the undersampling. The number of peaks at the same *intensity* as the central component (including the central component) is equal to the greatest common divisor of p and m , expressed as $\text{gcd}(p, m)$. Therefore, if the period size and the dimension length are not relatively prime, then the PSF power will have at least one perfectly aliased peak, resulting in $PSR = 1$. On the other hand, if the period size and the dimension length are relatively

prime, then the magnitude of the largest aliased peak in the PSF power will be diminished, resulting in $\text{PSR} > 1$. In summary

$$\begin{aligned} \gcd(p, m) > 1 &\rightarrow \text{PSR} = 1 \\ \gcd(p, m) = 1 &\rightarrow \text{PSR} > 1 \end{aligned} \quad (8)$$

The effect of this phenomena in 2D is demonstrated in Figure 9 for R_F , F and $\tilde{U}_F$ sampling on a grid with dimension length $m = 64$, where the coverage for R_F sampling is varied over the sequence $c \in \{1/2, 1/3, \dots, 1/20\}$ and the coverages for F and $\tilde{U}_F$ sampling are performed at the corresponding uniform undersampling periods $p \in \{2, 3, \dots, 20\}$. Fifty schedules are computed at each coverage value for R_F and $\tilde{U}_F$, as they include randomization, and only one schedule is constructed for F , which does not. The mean PSR values for R_F sampling (Figure 9A) follow a decay curve that scales with sample coverage. The PSR values for F sampling (Figure 9B) and mean PSR values for $\tilde{U}_F$ (Figure 9C), oscillate between a decay curve for when $\gcd(m, p) = 1$ and a nearly constant baseline for when $\gcd(m, p) > 1$.

We are now able to address the two unexpected findings in Figures 4 and 5, as noted at the beginning of this section. The $c = 20\%$ coverage corresponds to an undersampling period of $p = 5$, which has $\gcd(p, m) = 1$ for all of the dimension lengths in Table 1, except for $m = 40$. In contrast, the $c = 50\%$ coverage corresponds to an undersampling period of $p = 2$, which has $\gcd(p, m) > 1$ for all of the dimension lengths in Table 1.

The aliases induced by periodic undersampling in Figure 8 are disrupted by altering the relationship between undersampling period and dimension length. Similarly, Bretthorst [28, Figure 6] demonstrated that perfect aliases resulting from uniform on-grid sampling are eliminated by strategically introducing off-grid sampling points. This demonstrates that aliasing can be influenced by either disrupting the regularity of the data spacing (i.e. including off-grid points) or by ensuring that whatever regularity exists in the data spacing is not a divisor of the dimension length.

Conclusion

The PSR analysis of the set of carefully constructed sample schedules defined in Equation 3 elucidates connections between the reduction of NUS-induced aliasing and the dimensionality of the sample space. The approach is based on the PSF and independent of synthetic spectra analysis and specific reconstruction techniques – the PSR-based analysis presented here is a general and robust tool for characterizing and optimizing a broad range of NUS schemes.

In the first stage of analysis, individual data sets reveal an exponential relationship between dimensionality and PSR, with the strength of the exponent modulated by the fraction of indirect dimensions that include randomization. In the second stage of analysis, the comparison of PSR data sets uncovers an additional linear factor of dimensionality that partial-component NUS confers relative to full-component NUS by introducing randomization across the quadrature phases. Together, these two relationships define the

PSR expressions in Equation 7 and provide a powerful description of the importance of dimensionality, randomization and partial-component NUS.

The seemingly unconditional benefits of higher dimensionality demonstrated here utilized highly structured NUS schemes, designed to elicit exaggerated effects and ease visualization and characterization. The effects observed using more typical NUS schemes will be smaller than the effects shown here, but nonetheless may still be exploited. There are three additional factors involving the utilization of increased dimensionality that must be addressed. First, for fixed experiment time, additional dimensions come at the cost of shorter evolution times, resulting in a reduction in resolution. Nevertheless peaks unresolved along one dimension may be distinguishable along newly introduced dimensions. Second, higher dimensional experiments require longer pulse sequences, which cause signal loss due to relaxation prior to FID acquisition. However, the scaling relationships in Equation 7 indicate that the largest marginal improvements in PSR are attained by increasing dimensionality of low dimensional experiments, thereby avoiding excessive signal loss observed in the highest dimensional experiments. Alternatively, Kupce and Freeman [29] described a technique for constructing high-dimensional spectra from a series of conventional low dimensional experiments. Combining their insights with partial-component NUS offers another potential route to further improvements in multidimensional NMR spectra. Finally, it is clear that once the sampling artifacts are reduced to the level of experiment noise, there will be no additional benefits from increased dimensionality. If this regime is reached, the degrees-of-freedom afforded by partial-component NUS may be employed to shorten experiment time, rather than enhance PSR.

While the need for randomization in an NUS schedule has long been a guiding heuristic, the present analysis calls attention to a subtle synergistic link between randomization along the time dimensions and randomization along the quadrature phase “dimension” through partial-component NUS. The linear improvement that partial-component NUS brings relative to full-component NUS is only realized when the time dimensions also include randomization ($\tilde{U}_p$ and R_p). The use of partial-component NUS without randomization in the time dimensions ($\tilde{U}_p$) produces no benefit relative to uniform undersampling ($\tilde{U}_p$). Of principle importance is that any existing full-component NUS scheme can benefit from a reformulation as partial-component NUS. Experiment time may be decreased with such an approach or NUS-induced artifacts may be further reduced.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding from the US National Institutes of Health (R21-GM102366) is gratefully acknowledged.

References

1. Barna JCJ, Laue ED, Mayger MR, Skilling J, Worrall SJP. Exponential sampling, an alternative method for sampling in two-dimensional NMR experiments. *Journal of Magnetic Resonance*. 1987; 73(1):69–77.

2. Laue, Ernest D.; Skilling, John; Staunton, James; Sibisi, Sibusiso; Brereton, Richard G. Maximum entropy method in nuclear magnetic resonance spectroscopy. *Journal of Magnetic Resonance* (1969). 1985; 62:437–452.
3. Laue ED, Mayger MR, Skilling J, Staunton J. Reconstruction of phase-sensitive two-dimensional NMR spectra by maximum entropy. *Journal of Magnetic Resonance* (1969). 1986; 68 (1):14–29.
4. Ding, Keyang; Gronenborn, Angela M. Novel 2D triple-resonance NMR experiments for sequential resonance assignments of proteins. *Journal of Magnetic Resonance*. 2002; 156(2):262–268. [PubMed: 12165262]
5. Kup e, riks; Freeman, Ray. Projection-reconstruction of three-dimensional NMR spectra. *Journal of the American Chemical Society*. 2003; 125(46):13958–13959. [PubMed: 14611222]
6. Kim, Seho; Szyperski, Thomas. GFT NMR, a new approach to rapidly obtain precise high-dimensional NMR spectral information. *Journal of the American Chemical Society*. 2003; 125(5): 1385–1393. [PubMed: 12553842]
7. Kup e, riks; Freeman, Ray. Projection-reconstruction technique for speeding up multidimensional NMR spectroscopy. *Journal of the American Chemical Society*. May; 2004 126(20):6429–6440. [PubMed: 15149240]
8. Kup e, riks; Freeman, Ray. Fast multidimensional NMR: radial sampling of evolution space. *Journal of Magnetic Resonance*. 2005; 173(2):317–321. [PubMed: 15780924]
9. Coggins, Brian E.; Venters, Ronald A.; Zhou, Pei. Generalized reconstruction of n-D NMR spectra from multiple projections: application to the 5-D HACACONH spectrum of protein G B1 domain. *Journal of the American Chemical Society*. 2004; 126(4):1000–1001. [PubMed: 14746450]
10. Mobli, Mehdi; Stern, Alan S.; Hoch, Jeffrey C. Spectral reconstruction methods in fast NMR: reduced dimensionality, random sampling and maximum entropy. *Journal of Magnetic Resonance*. Sep; 2006 182(1):96–105. [PubMed: 16815055]
11. Maciejewski, Mark W.; Mobli, Mehdi; Schuyler, Adam D.; Stern, Alan S.; Hoch, Jeffrey C. Data sampling in multidimensional NMR: Fundamentals and strategies. *Topics in Current Chemistry*. 2012; 316:49–78. [PubMed: 21773916]
12. Mobli, Mehdi; Maciejewski, Mark W.; Schuyler, Adam D.; Stern, Alan S.; Hoch, Jeffrey C. Sparse sampling methods in multidimensional nmr. *Physical Chemistry Chemical Physics*. 2012; 14:10835–10843. [PubMed: 22481242]
13. Kazimierczuk, Krzysztof; Zawadzka, Anna; KoŹmi ski, Wiktor. Optimization of random time domain sampling in multidimensional NMR. *Journal of Magnetic Resonance*. 2008; 192(1): 123–130. [PubMed: 18308599]
14. Hoch, Jeffrey C.; Maciejewski, Mark W.; Filipovic, Blagoje. Randomization improves sparse sampling in multidimensional NMR. *Journal of Magnetic Resonance*. Aug; 2008 193(2):317–320. [PubMed: 18547850]
15. Hyberts, Sven G.; Takeuchi, Koh; Wagner, Gerhard. Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data. *Journal of the American Chemical Society*. 2010; 132(7):2145–2147. [PubMed: 20121194]
16. Marion, Dominique. Combining methods for speeding up multi-dimensional acquisition. sparse sampling and fast pulsing methods for unfolded proteins. *Journal of Magnetic Resonance*. 2010; 206(1):81–87. [PubMed: 20594882]
17. Schuyler, Adam D.; Maciejewski, Mark W.; Arthanari, Haribabu; Hoch, Jeffrey C. Knowledge-based nonuniform sampling in multidimensional NMR. *Journal of Biomolecular NMR*. 2011; 50(3): 247–262. [PubMed: 21626215]
18. Schmieder, Peter; Stern, Alan S.; Wagner, Gerhard; Hoch, Jeffrey C. Improved resolution in triple-resonance spectra by nonlinear sampling in the constant-time domain. *Journal of Biomolecular NMR*. 1994; 4(4):483–490. [PubMed: 8075537]
19. Lustig, Michael; Donoho, David; Pauly, John M. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*. 2007; 58(6):1182–1195. [PubMed: 17969013]
20. Coggins, Brian E.; Zhou, Pei. Polar fourier transforms of radially sampled NMR data. *Journal of Magnetic Resonance*. 2006; 182(1):84–95. [PubMed: 16820311]

21. Coggins, Brian E.; Zhou, Pei. Sampling of the NMR time domain along concentric rings. *Journal of Magnetic Resonance*. 2007; 184(2):207–21. [PubMed: 17070715]
22. Kazimierczuk, Krzysztof; Zawadzka, Anna; Kořmíski, Wiktor; Zhukov, Igor. Lineshapes and artifacts in multidimensional Fourier transform of arbitrary sampled NMR data sets. *Journal of Magnetic Resonance*. 2007; 188(2):344–356. [PubMed: 17822933]
23. Hoch, Jeffrey C.; Stern, Alan S.; Donoho, David L.; Johnstone, Iain M. Maximum entropy reconstruction of complex (phase-sensitive) spectra. *Journal of Magnetic Resonance*. 1990; 86: 236–246.
24. Donoho, David L.; Johnstone, Iain M.; Stern, Alan S.; Hoch, Jeffrey C. Does the maximum entropy method improve sensitivity? *Proceedings of the National Academy of Sciences of the United States of America*. Jul; 1990 87(13):5066–5068. [PubMed: 11607089]
25. States DJ, Haberkorn RA, Ruben DJ. A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants. *Journal of Magnetic Resonance* (1969). 1982; 48(2): 286–292.
26. Maciejewski, Mark W.; Fenwick, Matthew; Schuyler, Adam D.; Stern, Alan S.; Gorbatyuk, Vitaliy; Hoch, Jeffrey C. Random phase detection in multidimensional NMR. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108(40):16640–16644. [PubMed: 21949370]
27. Schuyler, Adam D.; Maciejewski, Mark W.; Stern, Alan S.; Hoch, Jeffrey C. Formalism for hypercomplex multidimensional NMR employing partial-component subsampling. *Journal of Magnetic Resonance*. 2013; 227:20–24. [PubMed: 23246651]
28. Larry Bretthorst G. Nonuniform sampling: Bandwidth and aliasing. *Concepts in Magnetic Resonance Part A*. 2008; 32(6):417–435.
29. Kupče, Eriks; Freeman, Ray. Hyperdimensional nmr spectroscopy. *Progress in Nuclear Magnetic Resonance Spectroscopy*. 2008; 52(1):22–30.

Highlights

- Nonuniform sampling of multidimensional NMR introduces artifacts.
- Randomizing offset of columns in a uniformly undersampled grid suppresses artifacts.
- Artifact suppression is exponentially related to dimensionality of sample space.
- Randomizing along “quadrature phase” yields synergistic improvement.

Figure 1. Full-Component Sample Schedules and PSFs

Sample schedules (top plot in each panel) are shown on a 64×64 grid with 12.5% coverage; black squares are sampled, white squares are not, the $t_1 = t_2 = 1$ time point is in the lower left corner of each plot. The power of the corresponding PSF is shown in the bottom plot in each panel. Full-component random sampling (**Panel A**) shows a central component with very low surrounding values in its PSF power (i.e. high PSR). At the other extreme, full-component constant-offset undersampling (**Panel B**) produces perfectly aliased peaks in its PSF power (i.e. PSR=1). Full-component random-offset undersampling (**Panel C**) introduces randomization of the undersampling offset along a *time* dimension, resulting in “squashed” artifacts distributed across f_2 and a PSR value approaching that of random sampling.

Figure 2. Partial-Component Sample Schedules and PSFs

Partial-component random sampling (**Panel A**), partial-component constant-offset undersampling (**Panel B**) and partial-component random-offset undersampling (**Panel C**) are shown in the same grid sizes, sample coverage and plot layouts as described in Figure 1. The partial-component sampling schemes depicted here employ randomization along the “quadrature phase dimension”, allowing independent sampling across each quadrature phase (i.e. “RR”, “RI”, “IR”, “II”), unlike the full-component schedules in Figure 1, which have the same sampling across all quadrature phases.

Figure 3. PSR for Random Sampling

PSR values for random sampling at full-component (**Panel A**) and partial-component (**Panel B**) according to the parameters in Table 1 are shown on log-log plots versus grid size with dimensionality (d) coded by color and coverage (c) coded by shape according to the legend. For each combination of dimensionality and coverage the data points are nearly collinear; the slopes of these data clusters correspond to the α parameter in Equation 6, and are shown in Figure 6. The standard deviation of each R_F and R_P data point is given in Supplemental Figures S1 and S2, respectively.

Figure 4. PSR for Constant-Offset Undersampling

PSR values for constant-offset undersampling at full-component (**Panel A**) and partial-component (**Panel B**) according to the parameters in Table 1 are shown on log-log plots versus grid size with dimensionality (d) coded by color and coverage (c) coded by shape according to the legend. All of the data points except those for $c = 20\%$ and $m = 40$ have $PSR=1$, indicative of the perfect aliases introduced by uniform undersampling. The remaining PSR values are slightly above one (note the scaling on the vertical axis); this deviation is explained in the “Periodicity of the DFT” section. The $\bar{U}_P$ data indicate that randomization along the “quadrature phase dimension” is not sufficient to reduce the sampling artifacts relative to $\bar{U}_F$. The standard deviation of each R_F and R_P data point is given in Supplemental Figures S3 and S4, respectively.

Figure 5. PSR for Random-Offset Undersampling

PSR values for random-offset undersampling at full-component (**Panel A**) and partial-component (**Panel B**) according to the parameters in Table 1 are shown on log-log plots versus grid size with dimensionality (d) coded by color and coverage (c) coded by shape according to the legend. For each combination of dimensionality and coverage the data points are nearly collinear; the slopes of these data clusters correspond to the α parameter in Equation 6, and are shown in Figure 6. The data follow increasing tiers, indicating that for a fixed grid size and fixed coverage, higher dimensional schedules perform better than lower dimensional ones. The standard deviation of each R_F and R_P data point is given in Supplemental Figures S5 and S6, respectively.

Figure 6. PSR Slope Values

The α values from Equation 6 are extracted from the PSR data in Figures 3–5 and shown here. The cluster of bars at each number of indirect dimensions is ordered left to right as listed in the legend from the top down. The α values for each schedule type correlate with the number of indirect dimensions that include randomization proportional to the total number of indirect dimensions. Constant-offset undersampling includes no randomization ($\alpha \approx 0$), random sampling includes randomization across all of its dimensions ($\alpha \approx 1$) and random-offset undersampling includes randomization across $d - 1$ of its d dimensions, yielding a $\frac{d-1}{d}$ progression marked by the gray horizontal lines extending to each cluster of bars. Additionally, the bars are marked with “F” for full-component and “P” for partial-component, highlighting the ability of the partial-component schemes R_P and $\tilde{U}_P$ to outperform their corresponding full-component schemes R_F and $\tilde{U}_F$, respectively.

Figure 7. Partial-Component NUS vs Full-Component NUS

The PSR data comparison for random sampling (**Panel A**) shows the data for R_P (Figure 3B) divided by the parameter matched values for R_F (Figure 3A). The PSR data comparison for random-offset undersampling (**Panel B**) shows the data for $\tilde{U}_P$ (Figure 5B) divided by the parameter matched values for $\tilde{U}_F$ (Figure 5A). Unlike the previous PSR figures, this figure shows PSR *ratios* on a linear vertical axis. The PSR ratios are all greater than one and increase linearly with dimensionality (as visually illustrated by the 3 parallel dashed lines in each panel), thus prompting the d scaling factor in Equation 7.

Figure 8. PSF Slices for f_1 Sampling at Various Coverages

The plots show one-dimensional PSF slices along f_1 , the undersampled dimension, of 2D f_1 sampling schedules. f_1 is of length $m = 64$ and the uniform undersampling is taken at period lengths $p = \{2, 5, 10, 20\}$. Each PSF trace has p peaks. The number of peaks with the same intensity as the central component is given by the greatest common divisor: $\gcd(p, m)$. Note that the PSF for $p = 2$ has only one aliased peak, but it is at the same intensity as the central component. In contrast, the PSF for $p = 5$ has more aliased peaks, but none are at the same

intensity as the central component. This observation is the reason for the PSR data at $c = 20\%$ performing better than at $c = 50\%$ for $\tilde{U}_F$ and $\tilde{U}_P$ sampling (Figure 5).

Figure 9. Periodicity of the DFT

R_F sampling (**Panel A**), F sampling (**Panel B**) and $\tilde{U}_F$ sampling (**Panel C**) are performed on a 64×64 grid. R_F sampling is performed at coverage values $c \in \{1/2, 1/3, \dots, 1/20\}$; F and $\tilde{U}_F$ sampling are performed at the corresponding coverage values specified by the uniform undersampling periods $p \in \{2, 3, \dots, 20\}$. Fifty schedules are computed at each coverage value for R_F and $\tilde{U}_F$ with each PSR value marked by a small circle and the mean PSR value at each coverage marked by a large circle. F sampling is performed once at each coverage value, as it does not include any randomization; its PSR values are marked with large circles. A solid line is used as a visual guide to connect the mean PSR values computed at sequential coverage values. The large circles are highlighted with gray for coverages of $\{50\%, 20\%, 10\%, 5\%\}$, which correspond to the sample coverage values used in the PSR analysis.

Table 1

Sample Schedule Parameters

The range of maximum increment values is chosen for each dimension to maximize the overlap in grid size, while avoiding unnecessary computation (i.e. 2D computations extend out to $m = 128$, but 5D stops at $m = 24$). Each combination of dimensionality (d), maximum increment (m) and sample coverage (c) is repeated 50 times for R_F , R_P , $\tilde{U}_F$ and $\tilde{U}_P$ sampling and only once for F , which is deterministic. p sampling is repeated 50 times for $c = 20\%$ and $m = 40$, for which $PSR = 1$, and only once at all other parameter combinations at which $PSR = 1$ (see the “Periodicity of the DFT” section for more information).

d	m									
	16	24	32	40	48	56	64	96	128	
2	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
5	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
<hr/>										
coverage										$c \in \{5\%, 10\%, 20\%, 50\%\}$