

Non-Uniform Sampling in NMR Spectroscopy and the Preservation of Spectral Knowledge in the Time and Frequency Domains

Manpreet Kaur,¹ Callie M. Lewis,¹ Aaron Chronister,^{1,2}
Gabriel S. Phun,^{1,3} and Leonard J. Mueller^{1,*}

¹Department of Chemistry, University of California, Riverside, 92521

²Current address: Department of Physics and Astronomy, University of California, Los Angeles, 90095

³Current address: Department of Chemistry, University of California, Irvine, 92697

* corresponding author: leonard.mueller@ucr.edu; ORCID ID: 0000-0002-2607-9875

Abstract

The increased sensitivity under weighted non-uniform sampling (NUS) is demonstrated and quantified using Monte Carlo simulations of nuclear magnetic resonance time- and frequency-domain signals. The concept of spectral knowledge is introduced and shown to be superior to the frequency-domain signal-to-noise ratio for assessing the quality of NMR data. Two methods for rigorously preserving spectral knowledge and the time-domain NUS knowledge enhancement upon transformation to the frequency domain are demonstrated, both theoretically and numerically. The first, non-uniform weighted sampling using consistent root-mean-square noise (NUWS-UCR), is applicable to data sampled on the Nyquist grid, while the second, the block Fourier Transform using consistent root-mean-square noise (BFT-UCR), can be used to transform time-domain data acquired with arbitrary, off-grid non-uniform sampling.

Key Words

non-uniform sampling, NUS, NMR, non-uniform weighted sampling, NUWS, block Fourier transform, Monte Carlo, spectral knowledge, spectral reconstruction

1. Introduction

A primary focus in the development of NMR spectroscopy has been placed on increasing sensitivity – the ability to detect weak signal peaks amongst the background noise. Experimental techniques for polarization transfer,¹ including hyperpolarization from DNP,^{2,3} parahydrogen,⁴ and photo-CIDNP,⁵ and hardware developments such as higher magnetic fields⁶ and cryoprobes⁷ are just a few examples of the many approaches taken to accomplish this. The technique of weighted non-uniform sampling (NUS) in the indirect dimensions has also been proposed to increase sensitivity⁸⁻¹² as well as shorten the acquisition of multidimensional spectra.¹³⁻¹⁸ Non-uniform sampling refers to any pattern of collecting samples that is done at irregular time intervals or with variable numbers of transients taken at different time points.¹⁹⁻²⁶ To achieve higher sensitivity with a finite number of measured points, weighted NUS places more samples at the early time in the free induction decay (FID) when the signal is stronger.⁸⁻¹² The predicted gains in sensitivity under weighted NUS (1.2 - 2)^{11,27} are not as large as some of the techniques mentioned above, but can be incorporated into any multidimensional experiment and compound with the number of indirect dimensions.²⁸

While the development of non-uniform sampling began over 30 years ago^{8,9} and numerous examples have been presented in the literature that demonstrate increased sensitivity,^{11,27-31} the approach is only slowly being embraced by the wider NMR community. Two salient features of uniformly-sampled time-domain data that contribute to its preeminent status are: (1) the ease with which it can be transformed into the frequency domain; and (2) the fundamental equivalence – guaranteed by virtue of the Fourier transform (FT) – of the time and frequency representations.^{32,33} Importantly, the FT is linear, so both signal and noise, as well as their ratio, can be equivalently defined and are preserved in both domains. In practice, this means that the frequency-domain spectrum is a lossless representation of the experimentally acquired data.

One challenge when working with data acquired under NUS is that direct Fourier transformation of the time-domain data to generate the frequency-domain spectrum is not always possible. A vast array of post-Fourier processing techniques such as maximum entropy reconstruction,^{34,35} forward Max Ent (FM),³⁶ maximum entropy interpolation (MINT),^{28,30} iterative soft thresholding (IST/RIST/hmsIST),^{37,38} SMILE,³⁹ NESTA,⁴⁰ FFT-CLEAN,⁴¹ SCRUB,⁴² DiffMap,⁴³ MDD,⁴⁴ and NUS-trained deep neural networks⁴⁵ have been developed, but lack the fundamental correspondence between the time and frequency domains enshrined by the Fourier transform. Most disconcerting, perhaps, is that the spectral reconstruction process is typically nonlinear, so the signal and the noise may not both be faithfully reproduced in the frequency domain. This often leads to the suppression of frequency-domain noise, and the meaning of the apparent signal-to-noise ratio (SNR) in this domain is obfuscated, making it difficult to judge the quality of the data or the fidelity of the reconstruction.⁴⁶ While many of the techniques mentioned above can be run in a quasi-linear regime, there is typically still some artificial suppression of the noise;³⁹ to our knowledge only MINT has been demonstrated to identically reproduce the frequency domain spectrum – both the signal and the noise – generated by the discrete FT (DFT) for a uniformly sampled FID.³⁰

Given the challenges in taking NUS data to the frequency domain, one option to assess the sensitivity gains of NUS has been to focus on the time-domain FID. The information in the frequency-domain spectrum cannot exceed that of the original time-domain data, and this leaves the pursuit of robust methods to construct the frequency domain without degrading the information content as a separate issue. The concept of the “intrinsic” signal-to-noise ratio (iSNR) of the time-domain data was introduced as a way to quantify the increased sensitivity of NUS data.^{27,47} The iSNR is an upper limit to the quality of the data that stands despite any increase in apparent SNR of the frequency-domain spectrum that results from post-acquisition manipulation (*e.g.*, apodization) or processing. With this metric for sensitivity, Palmer *et al.* proved the NUS Sensitivity Theorem, which for a decaying signal guarantees that the iSNR will improve for any sampling schedule that biases measurements to earlier times.²⁷

Here we take a complementary approach to evaluate the quality of time-domain NMR data acquired under uniform and non-uniform sampling that relies on spectral knowledge – a quantitative measure of the precision with which the spectral parameters of frequency offset, amplitude, and linewidth can be extracted from the experimental data.³³ As defined below, spectral knowledge describes the expected distribution of the fit parameters if the experiment were to be repeated multiple times; it therefore quantifies how well an extracted parameter is actually known. While spectral knowledge of the lineshape parameters is indeed proportional to the SNR, other factors contribute.³³ To compute spectral knowledge, we use the statistical approach of Monte Carlo simulation.⁴⁸ Monte Carlo methods are remarkably powerful, yet conceptually quite simple, and the approach closely follows the workflow of the actual experimental data acquisition and processing.

The advantage to using spectral knowledge over the standard frequency-domain SNR for quantifying the quality of NMR data is first demonstrated. This also serves to illustrate the fundamental equivalence of the two domains and points out several underappreciated aspects of signal and noise in the conjugate domains. Next, the advantage of non-uniform sampling in the indirect dimensions is quantified in the time domain for the case of an exponentially weighted sampling density. Notably, the predicted advantages of NUS in terms of sensitivity gains compared to uniform sampling (US) are to a large extent borne out. Two methods for reconstructing the frequency-domain data of densely ($\geq$ the Nyquist grid), non-uniformly sampled signals are then discussed: the direct Fourier transform of the sample-weighted FID acquired with non-uniform weighted sampling (NUWS) on the Nyquist grid,^{9,11,12} and the introduction of the “block Fourier Transform” (BFT) – an analytic Fourier Transform for time-domain signals acquired with non-uniform time-point sampling. Neither fully preserves the spectral knowledge available in the time domain. This motivates the introduction of two new methods – non-uniform weighted sampling using consistent RMS noise (NUWS-UCR) and the block FT using consistent RMS noise (BFT-UCR) – that are shown both theoretically and numerically to rigorously preserve spectral knowledge. NUWS-UCR and BFT-UCR are applicable to relatively densely sampled data and leave open the larger question of knowledge-preserving reconstruction for sparsely sampled NUS data. Here too we recommend the technique of Monte Carlo simulation as the gold standard in assessing the quality of any proposed acquisition scheme and the fidelity of the spectral reconstruction.

2. Theory and Methods:

2.1. Signal-to-Noise

In NMR, the signal-to-noise ratio (SNR) is defined as the quotient of the frequency-domain peak height, h , and the RMS noise, σ_N ,⁴⁹

$$SNR = \frac{h}{\sigma_N}. \quad (1)$$

The signal in the time domain is the digitized free-induction decay (FID): the collection of measured signal intensities at M specific time points

$$S_D(t) = \{(t_0, S(t_0)), (t_1, S(t_1)), \dots, (t_{M-1}, S(t_{M-1}))\}. \quad (2)$$

For uniformly sampled data in which $t_j = j\Delta$, and $t_{\max} = M\Delta$, the discrete Fourier transform gives the frequency-domain spectrum as the finite sum,

$$S_D^{DFT}\left(v = \frac{k}{M\Delta}\right) = \sum_j S(t_j) e^{-i2\pi k j / M}, \quad k = \left\{-\frac{M}{2}, -\frac{M}{2} + 1, \dots, 0, \dots, \frac{M}{2} - 1, \frac{M}{2}\right\} \quad (3)$$

which is typically evaluated at integer multiples, k , of the fundamental frequency, $1/(M\Delta)$.

Following Ernst,¹ for a time-domain signal that is a decaying complex exponential with amplitude A , frequency offset ν_o , and decay rate $R_2 = 1/T_2$,

$$S(t) = A e^{i2\pi\nu_o t} e^{-t/T_2}, \quad (4)$$

the signal-to-noise ratio can be approximated as

$$SNR = \frac{\sum_j S(t_j) e^{-i2\pi\nu_o t_j} \Big|_{\nu=\nu_o}}{\sigma_N} \approx \frac{\frac{M}{t_{\max}} \int_{t=0}^{t_{\max}} S(t) e^{-i2\pi\nu_o t} dt \Big|_{\nu=\nu_o}}{\sigma_N} = \frac{MAT_2 (1 - e^{-t_{\max}/T_2})}{t_{\max} \sigma_N}. \quad (5)$$

This expression applies equally to data acquired in both the direct and indirect dimensions of a multidimensional spectrum. It is important to note, however, that the noise characteristics in these domains are different. For the directly detected dimension, the noise is proportional to the square root of both the number of acquired points and the spectral width, F ,¹

$$\sigma_N \propto (MF)^{1/2} = M / \sqrt{t_{\max}} \quad (6)$$

In this case, the SNR can be written as

$$SNR(\text{direct}) \propto \frac{T_2 (1 - e^{-t_{\max}/T_2})}{\sqrt{t_{\max}}}, \quad (7)$$

which has a celebrated maximum at $t_{\max} = 1.26 T_2$.^{10,50} This means that every data point collected after $1.26T_2$ leads to a decrease in the SNR. As others have pointed out, the dependence on $t_{\max}$ is actually quite weak, and acquiring out to $3T_2$ leads to a decrease in the SNR of only 14%, while permitting significantly improved spectral resolution.⁵¹ It is also noted that in the direct dimension the SNR is independent of the dwell time and the number of samples, provided they are adjusted together for constant $t_{\max}$. This is a well-known result for anyone who has tried (and failed) to improve the experimental SNR by changing the spectral width at the spectrometer.

In the indirect dimensions, the noise is independent of the sampling rate and is proportional to the square root of the number of points acquired,

$$\sigma_N \propto \sqrt{M} \quad (8)$$

In this case,

$$SNR(\text{indirect}) \propto \frac{\sqrt{M} T_2 (1 - e^{-t_{\max}/T_2})}{t_{\max}}. \quad (9)$$

While it is tempting to simplify this expression further by presuming that $M \propto t_{\max}$, this should be resisted as Equation (9) only reduces to Equation (7) for constant dwell; Equation (9) is always true, however, regardless of the dwell time. It is standard to report the $SNR(\text{indirect})$ per root samples, as changing the number of samples in the indirect dimension under uniform sampling while keeping $t_{\max}$ constant simply adds to the signal-to-noise ratio stochastically (i.e., proportional to $\sqrt{M}$). Alternatively, one can focus on Equation (9) under conditions of constant M , which implies overall constant experimental acquisition time for a 2D NMR experiment. Equation (9) has a rather uncelebrated maximum at $t_{\max}=0$, which means that if the only consideration is maximum SNR, all samplings in the indirect dimension should be pushed to the earliest times when the signal is maximum. This shares some of the perspective of weighted non-uniform sampling strategies discussed below. Of course if the indirect FID is only sampled at early times all information regarding peak positions and linewidths would be lost and there would be no resolution of individual peaks; but the SNR would indeed be highest, a point that we return to below.

2.2. Spectral Knowledge

A complementary metric for the quality of NMR data is spectral knowledge (κ),³³ defined as the inverse of the normalized coefficient of variation (C_v) – the standard error (uncertainty), $\sigma(\mu)$, of an extracted spectral parameter relative to its expectation value, $\langle \mu \rangle$,

$$\kappa = \frac{1}{C_v} = \frac{\langle \mu \rangle}{\sigma(\mu)}. \quad (10)$$

The coefficient of variation is a measure of the dispersion of an extracted parameter from its ideally expected value. It answers the question that if an experiment were repeated, how close should an extracted parameter be to a previous measurement. We posit that spectral knowledge is the most relevant metric for a data set. While the normalized nature of C_v makes sense for parameters such as integrated peak amplitude and linewidth, it is somewhat misleading when comparing the precision of line positions at different frequency offsets, in which case $\sigma(\mu)$ is probably the better choice.

2.3. Monte Carlo Simulations of Spectral Knowledge

Both the spectral knowledge and the coefficient of variation require that the standard error, $\sigma(\mu)$, of an extracted parameter be determined. We accomplish this using Monte Carlo simulations,⁴⁸ which offer unbiased error estimates of spectral parameters in NMR signals, both as acquired in the time domain or as processed and displayed in the frequency domain. The approach is conceptually quite simple and is illustrated in Figure 1. First a model for the signal is taken as a decaying exponential with offset frequency ν_0 , amplitude A , and decay constant $R_2 = 1/T_2 = \pi w$, where w is the frequency-domain linewidth in Hz,

$$S(t) = A e^{i2\pi\nu_o t} e^{-\pi w t} . \quad (11)$$

The model is sampled at discrete time-domain points t_i (uniform or not) and combined with noise drawn from a suitable distribution (e.g., Gaussian white noise). Next, best estimates of the spectral parameters are obtained through a maximum likelihood/nonlinear-least-squares fitting by minimizing the chi-squared function

$$\chi^2(\theta) = \sum_{i=1}^N \sum_{k=1}^2 \left[\frac{S_{k,i} - S(k, t_i; \theta)}{\sigma_{n,i}} \right]^2 \quad (12)$$

with respect to the model parameters $\theta = \{A, \nu_o, w\}$.⁴⁸ $S_{k,i}$ is the experimentally-measured real ($k=1$) or imaginary ($k=2$) data point at time t_i , and $S(k, t_i; \theta)$ the corresponding value of the model for a given parameter set θ . The RMS noise in the time domain, σ_n , is explicitly allowed to depend on i , although it often does not.

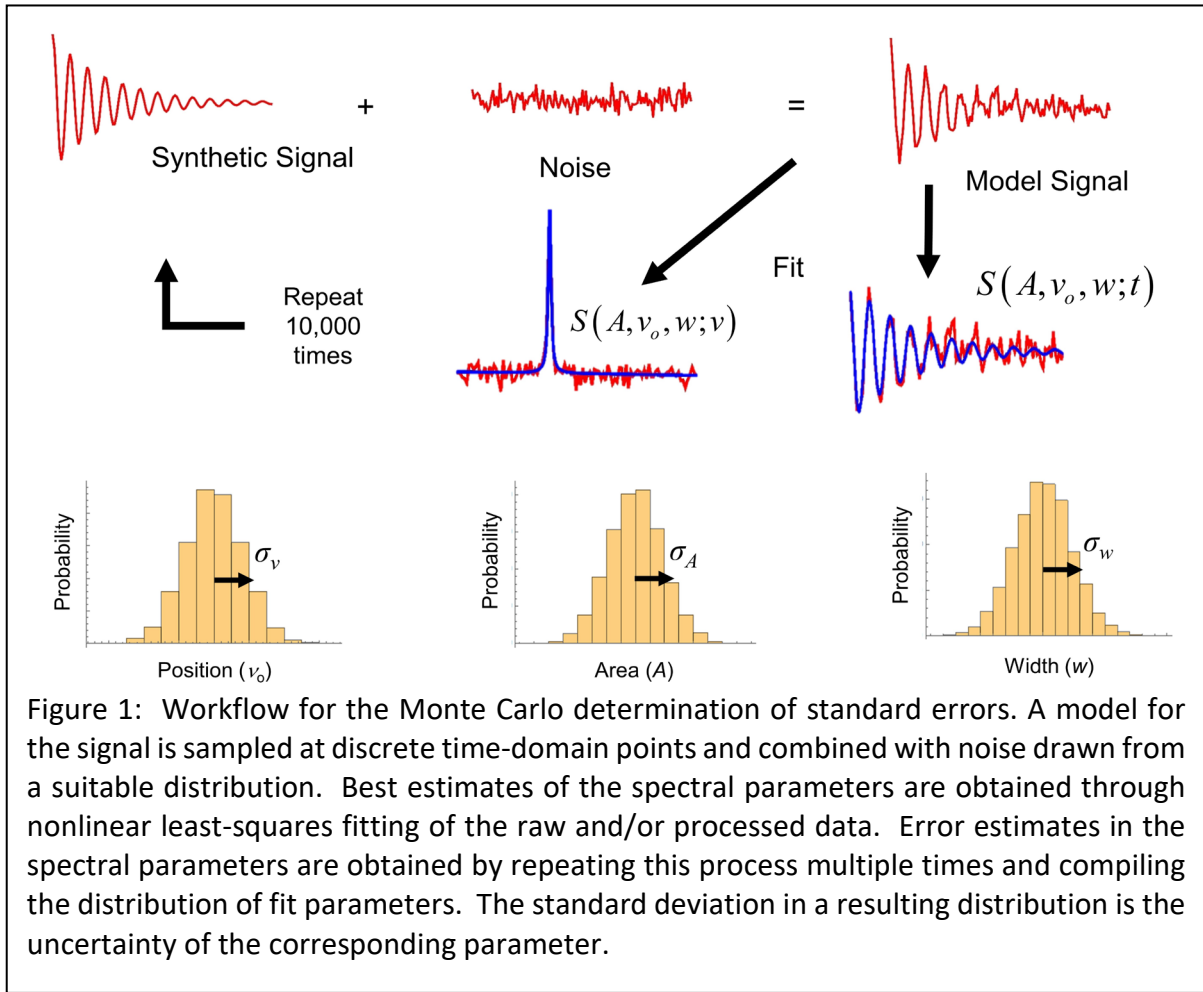
The time-domain data may also be transformed to the frequency domain and fit to the analytic transform of the model signal

$$\begin{aligned} S(\nu) &= \int_{t=0}^{t_{\max}} S(t) e^{-i2\pi\nu t} dt \\ &= \frac{A}{\pi w} \frac{1 - i2(\nu - \nu_o)/w}{1 + 4(\nu - \nu_o)^2/w^2} \left(1 - e^{-\pi w t_{\max}} e^{-i2\pi(\nu - \nu_o)t_{\max}} \right) \end{aligned} \quad (13)$$

where the finite acquisition time ($t_{\max}$) of the time-domain signal has been explicitly taken into account; in the limit that $t_{\max} \rightarrow \infty$ the Lorentzian lineshape is recovered. Recently, Dudley et al have derived additional analytical forms for the Fourier transform of time-truncated NMR signals with various apodization functions.⁵² In Equation (13),

$h = \frac{A(1 - e^{-\pi w t_{\max}})}{\pi w}$ can be identified as the frequency-domain peak height. Finally, error

estimates in the spectral parameters are obtained by repeating this process multiple times with new draws from the noise distribution and compiling the distribution of fit parameters; the standard deviation of this distribution is the uncertainty of the extracted parameter.



The advantages of the Monte Carlo approach are many. From a practical standpoint, it requires no advanced knowledge of statistical theory, and the results have a straightforward and obvious interpretation. It also corresponds closely to the experimental workflow: noise is introduced with well-known characteristics and minimal assumptions during the simulated signal acquisition, and no further assumptions need be made regarding the effects of processing on the noise. Monte Carlo analysis has been used previously to assess the efficacy of NUS acquisition and processing techniques, using peak detection probability as one of the metrics,^{39,53} and for demonstrating the increased precision from sharing parameters within and between multidimensional data sets.⁵² The related technique of IROC, which uses experimentally acquired time-domain data (including noise) combined with synthetic signals, has also been used to assess sensitivity and resolution metrics for various NUS acquisition probability distribution functions.⁵⁴ An alternate approach to determine the standard error is to use the Cramer-Rao Lower Bound (CRLB) derived from the Fisher information matrix.^{48,55} The CRLB is a powerful method for assessing statistical errors, but care must be taken in its direct application to processed data as post-acquisition manipulations may introduce correlated noise, which can substantially increase the complexity of the analysis.

2.4. Non-Uniform Weighted Sampling (NUWS)

NUWS is a variant of NUS that collects measurements on the uniformly spaced Nyquist grid, but with non-uniform numbers of repeated measurements ($n(t_k)$) at the time point t_k .⁹⁻

¹² The final collection of data points

$$S_D^{NUS-On}(t) = \left\{ \begin{array}{l} (t_0, S(t_0)_1), (t_0, S(t_0)_2), \dots, (t_0, S(t_0)_{n(t_0)}), \\ (t_1, S(t_1)_1), (t_1, S(t_1)_2), \dots, (t_1, S(t_1)_{n(t_1)}), \\ \dots, (t_{M-1}, S(t_{M-1})_{n(t_{M-1})}) \end{array} \right\} \quad (14)$$

is the on-grid time-domain NUS signal and can be fit directly in the time domain, but can also be averaged

$$\begin{aligned} S_D^{avg}(t) &= \left\{ \left(t_0, \frac{1}{n(t_0)} \sum_{k=1}^{n(t_0)} S(t_0)_k \right), \left(t_1, \frac{1}{n(t_1)} \sum_{k=1}^{n(t_1)} S(t_1)_k \right), \dots, \left(t_{M-1}, \frac{1}{n(t_{M-1})} \sum_{k=1}^{n(t_{M-1})} S(t_{M-1})_k \right) \right\} \\ &= \left\{ \left(t_0, \langle S(t_0) \rangle_{n(t_0)} \right), \left(t_1, \langle S(t_1) \rangle_{n(t_1)} \right), \dots, \left(t_{M-1}, \langle S(t_{M-1}) \rangle_{n(t_{M-1})} \right) \right\} \end{aligned} \quad (15)$$

or summed

$$S_D^{NUWS}(t) = \left\{ \left(t_0, \sum_{k=1}^{n(t_0)} S(t_0)_k \right), \left(t_1, \sum_{k=1}^{n(t_1)} S(t_1)_k \right), \dots, \left(t_{M-1}, \sum_{k=1}^{n(t_{M-1})} S(t_{M-1})_k \right) \right\} \quad (16a)$$

$$= \left\{ \begin{array}{l} \left(t_0, n(t_0) \langle S(t_0) \rangle_{n(t_0)} \right), \\ \left(t_1, n(t_1) \langle S(t_1) \rangle_{n(t_1)} \right), \\ \dots, \left(t_{M-1}, n(t_{M-1}) \langle S(t_{M-1}) \rangle_{n(t_{M-1})} \right) \end{array} \right\} \quad (16b)$$

at each point to give a time series that can be transformed to the frequency domain with the discrete Fourier Transform. The average FID (Equation (15)) is equivalent to what Simon and Köstler call apodization-weighted sampling,¹² while the sum process (Equation (16)) is referred to as non-uniform weighted sampling (NUWS).⁹ The construction of the NUWS and apodization-weighted time-domain signals from the NUS time-points is illustrated in Figure S1 of the Supporting Information.

For NUWS the underlying time-domain signal (absent noise) is the product of the signal, $S(t)$, and the sampling density, $n(t)$

$$S^{NUWS}(t) = S(t) n(t). \quad (17)$$

The frequency-domain lineshape under NUWS is therefore the convolution of the fundamental lineshape and the lineshape of the sampling density

$$S^{NUWS}(\nu) = S(\nu) \otimes n(\nu). \quad (18)$$

A simple procedure to generate the NUWS signal that avoids discontinuities and artefacts due to the integer restriction on $n(t_i)$ is to first calculate the average FID from the mean of the $n(t_i)$ data points at each time point and then multiply by the continuous sampling density function

$n(t)$. This is the procedure we will follow below for both time and frequency-domain analysis of NUWS.

2.5. The Block Fourier Transform

A second variation of NUS collects the FID with non-uniform spacing between the time-domain points. In this case, the number of transients collected at each point is constant, but the spacing is varied to satisfy the sampling density, $n(t)$. This collection of points is the off-grid time-domain NUS signal. The discrete FT of this signal presents multiple challenges. The DFT is functionally equivalent to treating the FID, $S_D(t)$, as a series of impulse (delta) functions weighted by the underlying signal, $S(t)$, at each measured time point t_j (Figure 2b)

$$S_D(t) = \sum_j S(t) \delta(t - t_j). \quad (19)$$

The analytic Fourier transform of this signal returns the discrete FT (Equation (3)), but can also be written as the convolution of the inherent frequency-domain lineshape, $S(\nu) = FT[S(t)]$, and the Fourier transform of the sampling schedule, $I(\nu) = FT\left[\sum_j \delta(t - t_j)\right]$,

$$S_D(\nu) = S(\nu) \otimes I(\nu). \quad (20)$$

It is the latter that leads to unacceptable baseline distortions that can obscure smaller peaks in the frequency domain when the time-domain points are sampled in a non-uniform manner.⁵⁶

Here we introduce a new approach – termed the “Block Fourier Transform” (BFT) – for the Fourier transform of discretely sampled data. In this method, the time-domain signal is considered to be a continuous function constructed as a series of linear (or potentially higher order polynomial) segments defined by the sampled points as shown in Figure 2c,

$$S_B(t) = \sum_j S_j^L(t); \quad (21)$$

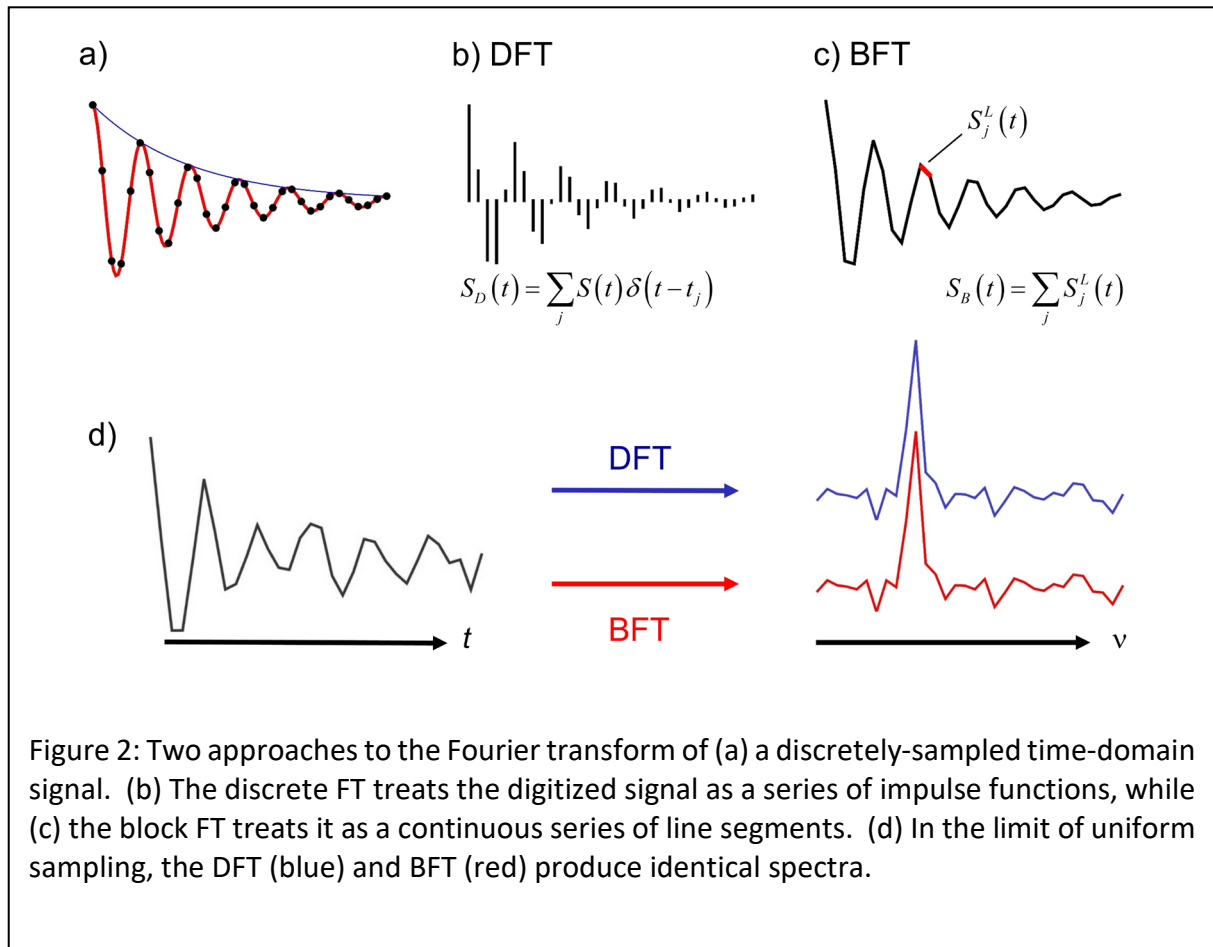
$S_j^L(\nu)$ is the line segment j joining measured point $(t_j, s(t_j))$ to point $(t_{j+1}, s(t_{j+1}))$. The Fourier transform of this sum can be written analytically as

$$S_B(\nu) = \sum_j S_j^L(\nu), \quad (22)$$

where $S_j^L(\nu)$ is the Fourier transformation of the j^{th} segment

$$S_j^L(\nu) = \begin{cases} \frac{\left(\frac{e^{-i2\pi\nu t_{j+1}} - e^{-i2\pi\nu t_j}}{4\pi^2\nu^2(t_{j+1} - t_j)} \right) (s(t_{j+1}) - s(t_j)) + i \frac{(e^{-i2\pi\nu t_{j+1}} s(t_{j+1}) - e^{-i2\pi\nu t_j} s(t_j))}{2\pi\nu}}{ } & \nu \neq 0 \\ \frac{1}{2} (s(t_{j+1}) + s(t_j)) (t_{j+1} - t_j) & \nu = 0 \end{cases}, \quad (23)$$

in which the value of the removable singularity at $\nu=0$ is explicitly given.



The major distinction between the discrete Fourier transform and the Block Fourier transform is how they approximate the integral transform. The discrete FT picks out individual points in the product of the signal and the integration kernel, while the block FT extrapolates the signal and integrates its product with the kernel over this block. The linear extrapolation and integration introduces a frequency-dependent attenuation. This attenuation applies equally to signal and noise, therefore not decreasing the local SNR, and can be removed by dividing out the filter response. For uniformly sampled data, the DFT is exactly reproduced in this process even in the presence of noise (Figure 2d).

The BFT can be applied to both uniformly and non-uniformly sampled data. This model assumes that the signal is bandwidth limited, *i.e.*, the signal frequency is less than half the sampling rate at all points – an assumption that can be relaxed under certain

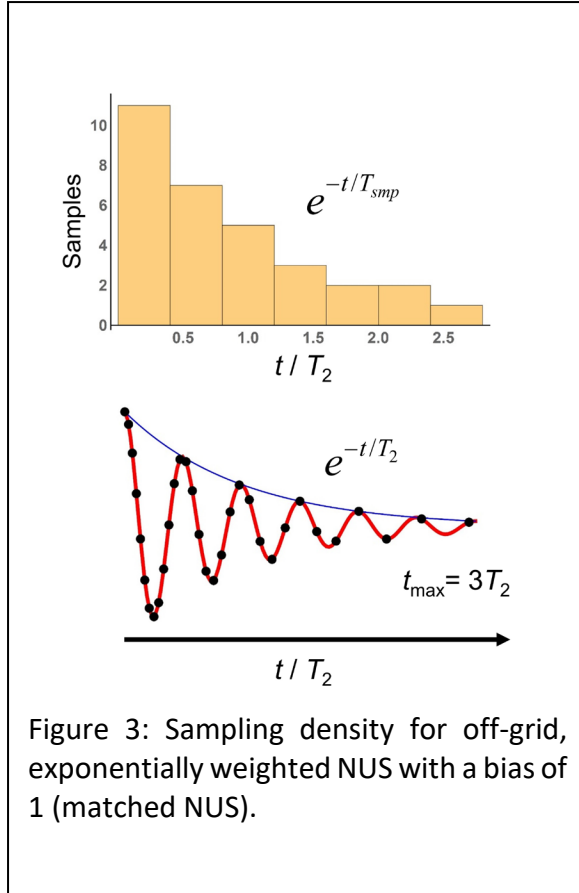
circumstances. Below we will make use of the BFT only on relatively densely sampled time-domain data in which case the filter response of the BFT is negligible. We note that the BFT is similar to extrapolation onto the Nyquist grid, but distinct as it can be directly applied to signals with arbitrary sample spacing.

2.6. Non-Uniform Sampling Strategies

A multitude of weighted and unweighted non-uniform sampling strategies have been introduced.^{8,22,24,56-65} Here, we focus on exponentially biased sampling in the dense-sampling regime, as it highlights the sensitivity advantages and trade-offs associated with weighted NUS.¹⁰ Exponentially biased NUS is characterized by a first-order sampling decay rate, $1/T_{\text{smp}}$, which parameterizes a sampling density

$$n(t) \propto e^{-t/T_{\text{smp}}}. \quad (24)$$

The sampling decay rate is often reported in terms of a bias, $b = T_2/T_{\text{smp}}$, relative to the signal decay rate.¹⁰ A bias of 0 equates to uniform sampling, while increasing biases place greater numbers of measurements at earlier times in the FID. For on-grid NUS, this corresponds to more independent measures at the earlier Nyquist-grid time-points, while for off-grid NUS this corresponds to a faster pace of sampling earlier in the FID. Figure 3 shows the sampling distributions for off-grid exponentially weighted NUS with a bias of 1 (matched NUS).



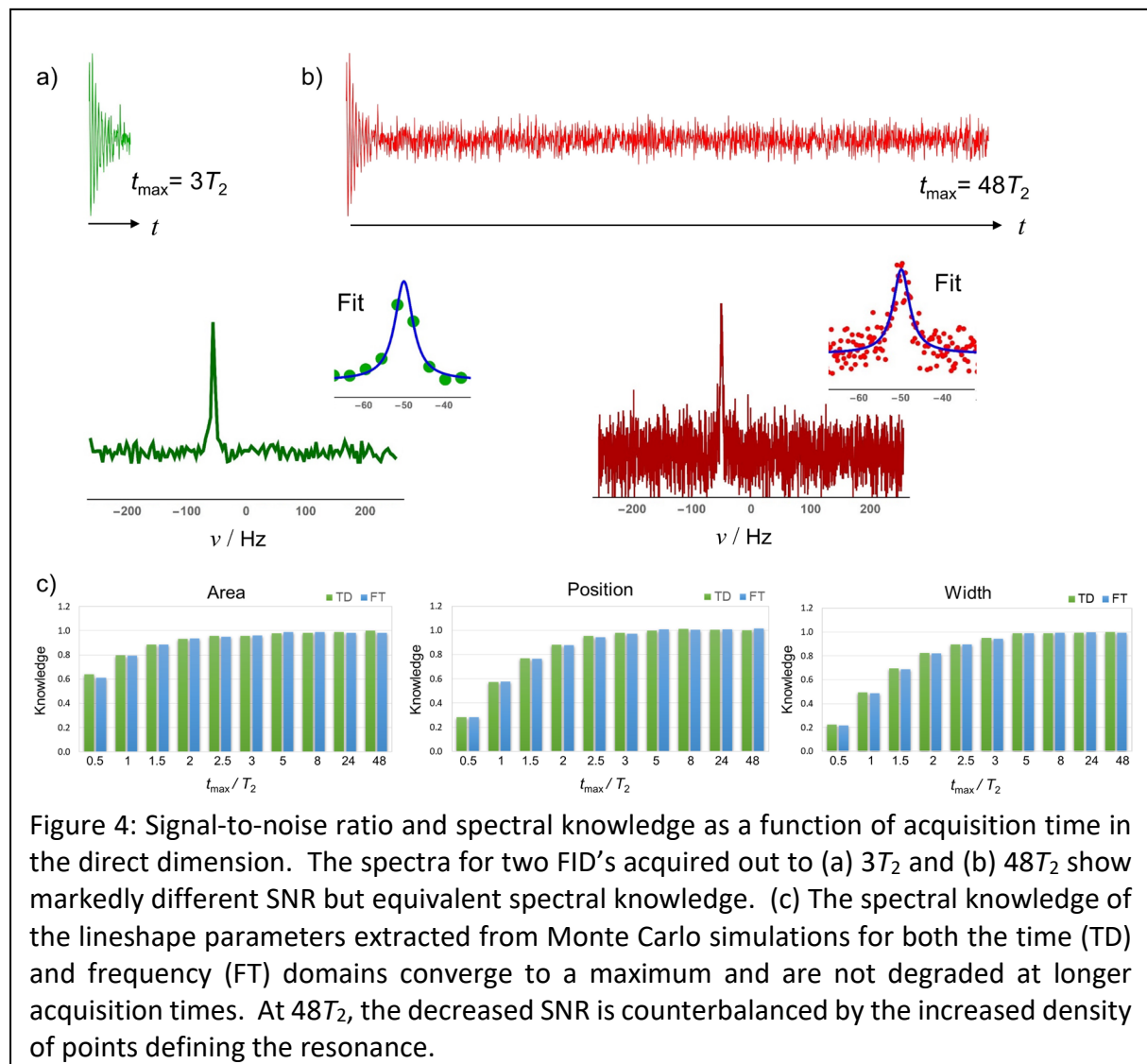
2.7. Numerical Procedures

Monte Carlo simulations were performed in the Mathematica programming environment.⁶⁶ For on-grid NUS, a time-domain FID was generated from 1024 complex-valued samples of the model signal on a Nyquist grid consisting of 256 time-domain points from $t=0$ to variable $t_{\max}$. The number of independent samples collected at each time point was chosen to conform to exponentially-weighted distributions of varying biases. The signal was modeled with amplitude $A=1$, frequency $\nu_o = 50$ Hz, and decay rate $R_2 = 1/T_2 = \pi w$ for $w=5$ Hz. For off-grid NUS, 1024 signal samples with exponentially increasing spacing were arranged as independent time-domain points from $t=0$ to $t_{\max}$; the spacing of the final two points remained within a factor of two of that for the uniformly sampled grid. Complex-valued noise was added to each of 1024 independent signal samples. The noise for each channel was chosen from a Gaussian distribution characterized by a mean of 0 and a standard deviation of 0.3. For uniform sampling and $t_{\max}=3T_2$, this produced a frequency-domain spectrum with a SNR of $\sim 30:1$. Nonlinear least-squares fitting was performed using the built-in Mathematica function “NonlinearModelFit.” Convergence of the Monte Carlo distribution was confirmed by visual inspection of the mean and standard deviations of the fit distributions as a function of the number of Monte Carlo cycles; a representative example is shown in Figure S2 (SI).

3. Results and Discussion

3.1. The Shortcomings of SNR: Length of Data Acquisition in the Direct Dimension

As a first example, we consider the effect of the length of data acquisition in the direct dimension on the precision with which the lineshape parameters of amplitude, frequency offset, and linewidth can be extracted in both the time and frequency domains. In the directly detected dimension, there is typically no overall experimental time penalty for acquiring longer signals, although as already noted longer acquisitions ($> 1.26 T_2$) lead to decreased SNR in the frequency domain due to the additional noise that is acquired.²⁷ This is illustrated in Figures 4a and b, which show the frequency-domain spectra generated from two otherwise identical time-domain FIDS (equivalent signal amplitude, decay rate, frequency offset, acquisition dwell time, and RMS time-domain noise) but with the FID in 4b acquired 16 times longer than that in 4a. There are clear differences in the SNR of the resulting spectra, and it would seem intuitive that the spectral parameters extracted from 4b should be less precise. This turns out to be incorrect.



To quantify the spectral knowledge in both the time and frequency domains as a function of acquisition time, a Monte Carlo approach was used in which a time-domain signal was generated with varying length from $0.5 T_2$ to $48 T_2$, but otherwise equivalent signal and RMS noise properties. In these simulations, the acquisition dwell time was kept constant, so the longer sampling times correspond to proportionately more data points being acquired (M in Equation (5)). These signals were then fit directly in the time domain to Equation (11) and also Fourier transformed and fit in the frequency domain to the truncated Lorentzian lineshape, Equation (13). At each acquisition time, the process was repeated 10,000 times with the same spectral parameters but new draws from the noise distribution, and the fit parameters tabulated. The mean and standard deviation of the resulting distributions are reported as the spectral knowledge (Equation (10)).

Figure 4c displays the spectral knowledge for the three lineshape parameters (amplitude, width, and offset) extracted from the time and frequency-domain fits as a function of the acquisition time. Note two important features of these results. First, knowledge is equal in the time and frequency domains; this is consistent with the fact that the Fourier transform is linear and produces a fully equivalent representation of the time-domain data. An important caveat is that the correct functional form for the frequency-domain signal – the time truncated Lorentzian (Equation (13)) – must be used in order to observe this correspondence; fits to a pure Lorentzian lineshape were found to have larger associated errors in the frequency domain, particularly for short acquisition times. Second, knowledge converges to a maximum and is not degraded at longer acquisition times. This is surprising given the obvious differences in the spectra of Figures 4a and b. However, this convergence is actually quite intuitive from the perspective of the time domain, as the additional points fall where the model signal has decayed and are simply distributed about zero. In this sense, the noise of these latter points cannot adversely influence the extracted parameters – those are determined where there is signal. The situation from the perspective of the frequency domain is only slightly more challenging to understand. Here, the increased acquisition time leads to a proportionate increase in the digital resolution, so for the example in Figure 4b there are 16 times the number of points defining the curve in the frequency domain compared to that in Figure 4a. At the same time, the RMS noise at each frequency-domain point is also increased, but only by the square-root of the increased points (Equation (6)), so a factor of 4 in this example. The additional 16 points combine in the same way as signal-to-noise ratios in NMR experiments, with the knowledge increasing as the square root of the number of points divided by the noise. Thus, knowledge is preserved in the frequency domain as well.

This example points out three important properties of NMR signals. First, the knowledge content of the frequency-domain spectrum is not defined solely by the SNR – a decreased SNR may be compensated for by increased digital resolution (i.e., increased density of data points).³³ Second, provided the correct corresponding functional forms are used (i.e., the time-truncated Lorentzian) and that both the real and imaginary data points are fit, spectral knowledge in the time and frequency domains is equivalent; this, of course, is no surprise and is guaranteed by the properties of the Fourier Transform.³² Finally, not all noise in the time domain is equivalent: noise that occurs after the signal has decayed has significantly decreased consequence, even when ultimately looking at the frequency-domain data.

3.2. Data Acquisition in the Indirect Dimension and NUS

In the indirect dimension, each additional point in the FID adds linearly to the overall experiment time. A relevant question to ask, then, is how a given number of samples should be arranged for optimum parameter knowledge. Options include uniform sampling out to some cutoff time (e.g., $1.5 T_2$, $3T_2$) and NUS with variously biased distributions and acquisition times. As noted above, we focus on exponential sampling with differing biases and maximum acquisition times, as it highlights the advantages and trade-offs associated with weighted NUS.

3.2.1. Single Peak

We begin by considering the effect of weighted NUS on intrinsic spectral knowledge for a time-domain signal consisting of a single resonance. Here, “intrinsic” designates the inherent spectral knowledge of the time-domain data before any post-acquisition manipulation or processing, in analogy to the iSNR.⁴⁷ The Monte Carlo approach was used to generate time-domain signals with Gaussian white noise, in which the total number of sampled time-domain points was held constant (*i.e.*, constant total experiment time) and the sampling-bias and maximum acquisition time varied. The resulting synthetic data sets were fit to the decaying complex exponential in Equation (11), and once more the distribution of fit parameters was used to define the spectral knowledge for each combination of sampling bias and maximum acquisition time. These parameters are reported as the NUS knowledge enhancement, the normalized ratio of the parameter knowledge under NUS to that for uniform sampling out to $t_{\max}=3T_2$,

$$\eta = \frac{\kappa(\text{NUS})}{\kappa(\text{US}; t_{\max}=3T_2)}, \quad (25)$$

and displayed in Figure 5. The NUS knowledge enhancement is a quantitative measure of the advantage to NUS acquisition relative to standard US. Larger NUS knowledge enhancements indicates a benefit in terms of the inherent knowledge (lower uncertainty) of a spectral parameter contained within the data. Note that by definition the NUS knowledge enhancement for US (bias=0) at $t_{\max}=3T_2$ is 1 for each of the fit parameters.

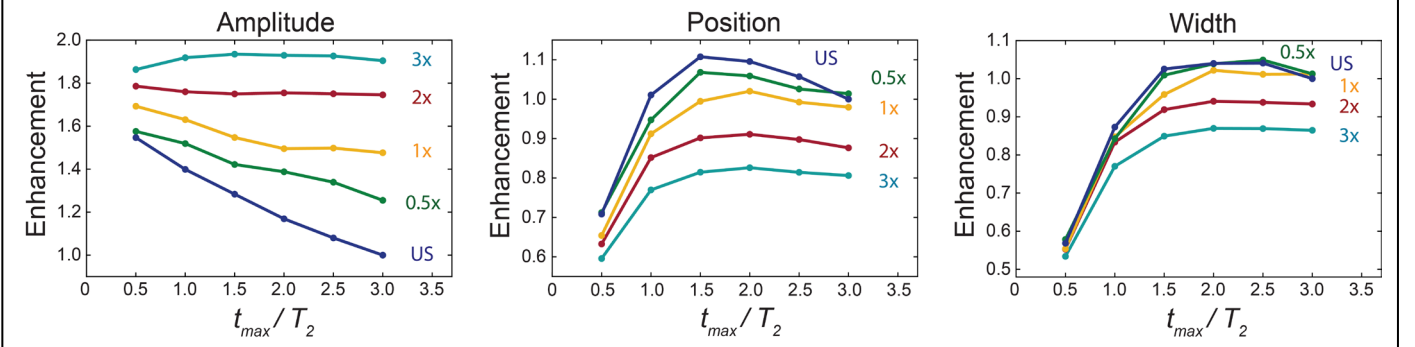


Figure 5: Intrinsic spectral knowledge reported as NUS knowledge enhancements for a single resonance under the indicated exponentially biased NUS strategies and varying $t_{\max}$. For each combination, the total number of time-domain points sampled was held constant, corresponding to equivalent total experiment times. Weighted NUS leads to increased spectral knowledge for peak amplitude at all biases and $t_{\max}$ and is therefore more sensitive than US. There is a concomitant decrease in spectral knowledge for peak position and linewidth under NUS. In all cases, the intrinsic spectral knowledge was extracted from the time-domain FID's.

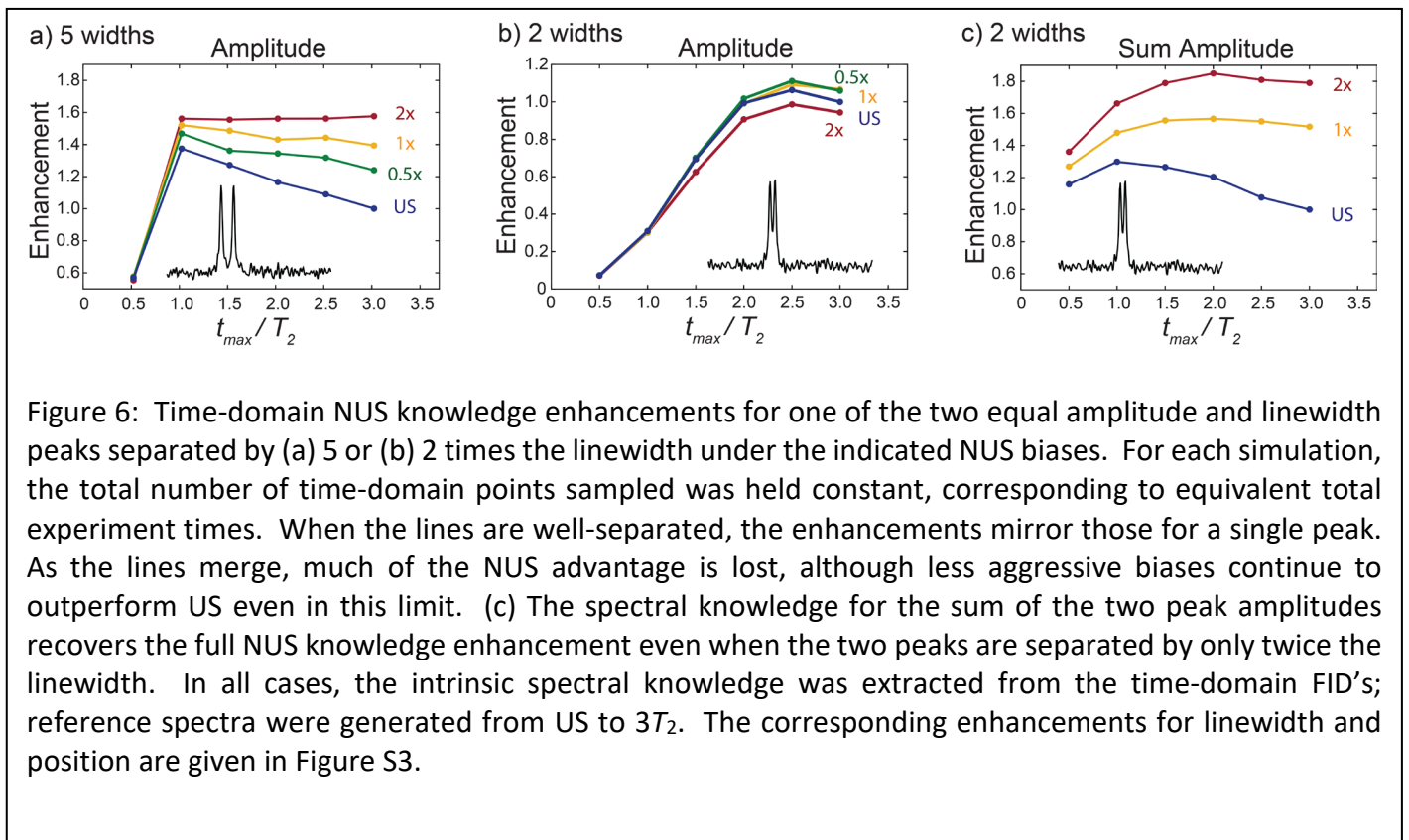
Perhaps the most immediate parameter of interest is the signal amplitude, which corresponds to the peak area in the frequency-domain and is directly proportional to the number of nuclear spins. Spectral knowledge of this parameter quantifies with what certainty a peak can be identified. Here, one of the most important predicted advantages of NUS is borne out: for all biases, NUS provides enhanced intrinsic spectral knowledge of peak amplitude, and is therefore more sensitive than US. The gains are 1.2-1.9 depending on bias and $t_{\max}$, but come at no change in the overall experiment time. Note that in all sampling schemes, the most sensitive approach for a finite number of samples is to acquire them at earlier times (shorter $t_{\max}$). This corresponds to the maximum noted for the SNR in Equation (9). This is not a practical strategy in most cases, however, as knowledge of linewidth and frequency offset require the signal at later times, as shown in Figure 5. But sample times $t_{\max} > 1.5T_2$ and biasing toward the earlier times with exponential sampling is a reasonable compromise, with matched NUS (bias=1) offering both good sensitivity increases relative to US with little degradation of knowledge in position and width relative to the uniformly sampled signal for $t_{\max} = 3T_2$.

The NUS enhancements for spectral knowledge track those for iSNR,²⁷ but it is important to recognize that spectral knowledge and iSNR are unique metrics and show a different dependence on $t_{\max}$. Notably, the spectral knowledge for peak amplitude, which plays a similar role to the SNR as a measure of the ability to detect a peak, converges with increasing $t_{\max}$; this is evident for larger biases in Figure 5, but is expected to occur for all exponential biases. This contrasts with the iSNR, which is a monotonically increasing function of $t_{\max}$.²⁷ The difference can be attributed in part to the fact that we have defined the NUS knowledge enhancement relative to a fixed reference point, uniform sampling to $3T_2$. That the NUS knowledge enhancements for all three properties converge is a consequence of the exponentially-biased sampling density, which itself converges with respect to $t_{\max}$. Because of the exponential decay, at some point a nominal increase in $t_{\max}$ makes no appreciable change to the distribution of measured points, and the spectral characteristics become fixed.

3.2.2. Two Peaks

When multiple peaks are present, there is a trade-off between the sensitivity advantage of acquiring at shorter time and the time-development necessary to resolve the signals into individual resonances. For two peaks, the spectral knowledge will depend on the relative ratio of the peak amplitudes and linewidths as well as their frequency separation. Here we consider a limited range of parameters in which the amplitude and linewidths vary by a factor of 2 – a not so uncommon situation in many 2D NMR experiments. The Monte Carlo procedure follows that outlined above, but with a signal that is composed of the sum of two complex, decaying exponentials. Two separate questions are addressed: first, the spectral knowledge of an individual peak within the pair; second, the spectral knowledge of the sum of the two peak amplitudes.

Figures 6a and 6b show the spectral knowledge for an individual peak within the pair, reported as the NUS knowledge enhancement, under various exponential NUS biases. In this case, the two peaks have equal amplitudes and linewidths and are separated by either 5 or 2 times the linewidth. When the peaks are well-separated, the NUS knowledge enhancements for the individual components mirror those of a single peak, and a choice can be made to forgo some knowledge of the peak position and linewidth relative to US in order to increase knowledge in the peak amplitude by 40%-60%. As the peaks get closer together, however, much of the NUS knowledge enhancement is lost. When the peaks are separated by twice the linewidth, the knowledge of the amplitude is only slightly better (10%) for matched exponential NUS compared to uniform sampling for $t_{\max}=3T_2$. This can be rationalized based on the need to sufficiently evolve the signals in order to resolve the lines and assign signal



amplitude appropriately to each. While this loss in sensitivity gain may temper enthusiasm for NUS, it is important to note that matched exponential NUS does not underperform relative to US, even in this limit.

A separate and very important feature of the data is revealed when the spectral knowledge of the sum intensity for the two peaks is considered. In the Monte Carlo simulations above, the sum intensity for the two partially resolved lines was also tracked at each pass of the Monte Carlo cycle by adding the amplitudes of the two components together. These are summarized in Figure 6c and show similar enhancements (factor of 1.5-1.8 for $t_{\max}=3T_2$) to those for a single peak under weighted NUS. The spectral knowledge of the overall intensity is greater under NUS than US even as the peaks begin to overlap; this is despite the fact that it is challenging to assign amplitudes to the individual resonances. The NUS spectral knowledge enhancement for the sum intensity is in most cases the more relevant choice when considering questions of sensitivity. Although the sorting of intensity into the separate peaks poses a challenge to NUS (as well as to US), the overall sensitivity still benefits considerably from non-uniform sampling.

Figure S4 shows the corresponding spectral knowledge under various exponential NUS strategies for the case of two peaks of unequal linewidth. The results are qualitatively similar to those above, noting that the effective sampling bias changes for the second peak as its linewidth changes. Again, matched NUS (defined by the linewidth of the first peak) remains a good compromise between increased knowledge in peak amplitude and decreased knowledge in position and linewidth. Once more, as the peaks become less well-resolved the advantages of NUS over US are maintained but decrease in size.

The implication of these simulations is that NUS is a robust method for increasing sensitivity under most experimental conditions. However, there is no guarantee that NUS will always outperform US when the goal is to assign intensity to individual peaks: Figure 6b shows a counterexample in which for two closely spaced lines, US outperforms NUS at high sampling bias. However, under most conditions NUS will be superior.

3.3. NUS and the Frequency Domain

There is considerable benefit – quantified in terms of intrinsic spectral knowledge of peak amplitude extracted in the time domain – to NUS in the indirect dimension(s) of NMR experiments. The analysis of NMR data, however, is typically performed in the frequency domain – and with good reason, as the interpretation of NMR data is tied to transition frequencies. While the same information is encoded in both domains, it is localized in the frequency domain and spread out in the time domain. Generating the frequency-domain spectrum from the time-domain FID is straightforward for US using the discrete Fourier transform. It is more complicated for NUS sampled data, and considerable effort is currently underway to develop robust methods to accomplish this.^{28,30,36-45} Here we consider two approaches that are applicable to dense, non-uniformly-sampled time-domain data: direct Fourier transform of the NUWS FID acquired on the Nyquist grid, and the block Fourier transform of NUS data acquired off the Nyquist grid. We demonstrate how with appropriately modified versions of each the gains in intrinsic spectral knowledge realized with NUS in the

time domain can be rigorously retained in the construction of the frequency-domain spectrum.

3.3.1. On-Grid Methods:

3.3.1.1. Non-Uniform Weighted Sampling

NUWS was one of the original methods proposed for acquiring and processing non-uniformly sampled data.⁹ NUWS collects variable numbers of transients at each time point and simply adds them together. NUWS has the advantage that the data remain on the Nyquist grid, allowing for the discrete FT of the resulting NUWS FID. This approach cannot be used to decrease the minimum number of time points acquired in the indirect dimension in order to shorten overall experimental time, but is potentially appropriate for sensitivity-limited experiments, such as natural abundance correlation. NUWS distorts the inherent lineshape, which in the frequency domain becomes the convolution of the underlying (truncated) Lorentzian and the FT of the weighting function. For the exponentially-weighted sampling densities considered here, this increases the Lorentzian linewidth based on the sampling bias.

Monte Carlo simulations were performed to determine how well the NUWS scheme preserves the NUS knowledge enhancements demonstrated above for time-domain data. To begin, a list of $n(t_1)$ samples of the model signal plus Gaussian white noise was generated at each time point on the Fourier grid (Equation (14)). Here, $n(t)$ conformed to an exponential distribution with a bias of 1 (matched NUWS). The data were then summed at each time point (Equation (16)), fit as the sum in the time domain, and Fourier transformed and fit in the frequency domain. The process was repeated 10,000 times and the knowledge of the fit parameters extracted from the resulting distributions. The known linewidth contribution due to the sampling bias was removed from the fit value so that direct comparisons between different biases could be made. The intrinsic spectral knowledge of the time-domain NUS data was also determined from the original data list before summation, which included multiple measurements at each time point, by direct fitting in the time domain; this corresponds to the 1x bias NUS curve in Figure 5. For computing the NUS knowledge enhancement, a reference US signal with the same total number of measurements was generated and analyzed in the time domain. These data are summarized in Figure 7a for a single peak with varying $t_{\max}$. As fitting the summed NUWS signal in the time domain and its subsequent FT and fitting in the frequency domain were found to be equivalent, only the results from the frequency-domain fits are shown.

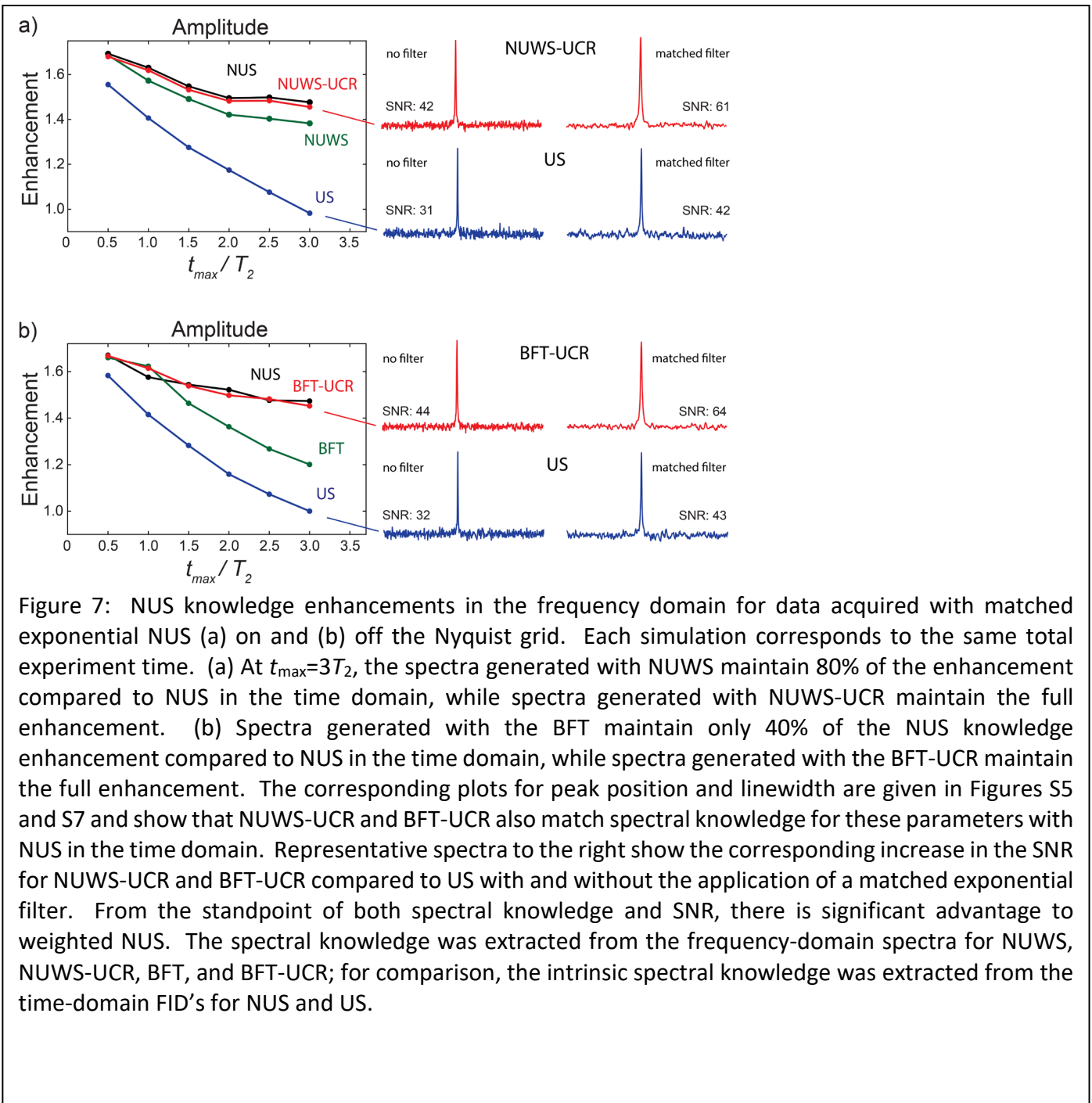


Figure 7: NUS knowledge enhancements in the frequency domain for data acquired with matched exponential NUS (a) on and (b) off the Nyquist grid. Each simulation corresponds to the same total experiment time. (a) At $t_{\max}=3T_2$, the spectra generated with NUWS maintain 80% of the enhancement compared to NUS in the time domain, while spectra generated with NUWS-UCR maintain the full enhancement. (b) Spectra generated with the BFT maintain only 40% of the NUS knowledge enhancement compared to NUS in the time domain, while spectra generated with the BFT-UCR maintain the full enhancement. The corresponding plots for peak position and linewidth are given in Figures S5 and S7 and show that NUWS-UCR and BFT-UCR also match spectral knowledge for these parameters with NUS in the time domain. Representative spectra to the right show the corresponding increase in the SNR for NUWS-UCR and BFT-UCR compared to US with and without the application of a matched exponential filter. From the standpoint of both spectral knowledge and SNR, there is significant advantage to weighted NUS. The spectral knowledge was extracted from the frequency-domain spectra for NUWS, NUWS-UCR, BFT, and BFT-UCR; for comparison, the intrinsic spectral knowledge was extracted from the time-domain FID's for NUS and US.

Figure 7a shows that NUWS realizes some, but not all, of the potential time-domain NUS knowledge enhancement with respect to peak amplitude. For matched bias and sampling to $3T_2$, the intrinsic NUS knowledge enhancement is 1.48, while NUWS returns 1.38. NUWS retains only 80% of the full NUS knowledge enhancement.

The reason NUWS underperforms can be understood by comparing the NUWS procedure to that for NUS in the time domain. NUWS sums the independent measures at each time point and then a nonlinear least-squares fit to the resulting NUWS FID is performed. In contrast, NUS keeps the replicate data as independent measures in the fitting procedures; that multiple time points fall at the same t_1 values is no complication to either the definition of chi-squared or the maximum likelihood fitting. The least-squares fitting for both the NUWS

and NUS data in the time domain is unweighted, which means that the estimate of the noise at each time point in the chi-squared function is assumed to be equal. But by construction of the NUWS signal, the noise at each time point is not equal, it is proportional to the square-root of the number of measurements that comprise the sum. When this is properly accounted for by setting $\sigma_{n,i} \rightarrow \sqrt{n(t_i)}\sigma_{n,i}$ in Equation (12) (i.e., using weighted least-squares fitting), the NUS knowledge enhancement can be fully recovered for the NUWS time-domain signal. However, there is no direct way to transfer the time-domain weightings to a frequency-domain chi-squared function, because each individual point in the time domain contributes to all frequencies. So, although the NUS knowledge enhancement can be realized by proper analysis of the NUWS time-domain FID, there is no obvious weighted least-squares approach to accomplish this in the frequency domain.

Monte Carlo simulations demonstrate that the time-domain signal constructed with apodization weighting (Equation (15)) also underperforms compared to NUS in the time domain (Figure S5). The failure of apodization weighting to achieve the full NUS enhancement can again be ascribed to the construction of a signal that has a variation in the noise across the time-domain points. While appropriately weighted least-squares fitting can once more recover the full NUS enhancement in the time domain, there is no simple remedy in the frequency domain.

3.3.1.2. Non-Uniform Weighted Sampling Using Consistent RMS Noise

The above analysis might seem an essential limitation to achieving the full NUS knowledge enhancement in the frequency domain. Fortunately, it also points to a straightforward solution: if the data can be acquired OR combined to have equal uncertainty (RMS noise) at each point in the time domain, then the frequency-domain spectrum produced by the DFT will fully preserve the time-domain spectral knowledge. Here we introduce a new approach to NUWS that achieves this by constructing the root-weighted signal

$$\begin{aligned}
 S_D^{NUWS-UCR}(t) &= \left\{ \left(t_0, \frac{1}{\sqrt{n(t_0)}} \sum_{k=1}^{n(t_0)} S(t_0)_k \right), \left(t_1, \frac{1}{\sqrt{n(t_1)}} \sum_{k=1}^{n(t_1)} S(t_1)_k \right), \dots, \left(t_{M-1}, \frac{1}{\sqrt{n(t_{M-1})}} \sum_{k=1}^{n(t_{M-1})} S(t_{M-1})_k \right) \right\} \\
 &= \left\{ \left(t_0, \sqrt{n(t_0)} \langle S(t_0) \rangle_{n(t_0)} \right), \left(t_1, \sqrt{n(t_1)} \langle S(t_1) \rangle_{n(t_1)} \right), \dots, \left(t_{M-1}, \sqrt{n(t_{M-1})} \langle S(t_{M-1}) \rangle_{n(t_{M-1})} \right) \right\}
 \end{aligned}
 \tag{26}$$

This produces a FID with equivalent noise characteristics at each time-domain point, so there is no longer an issue with varying weights (they are now rigorously the same) or how to handle them in the conjugate domain. We call this procedure *non-uniform weighted sampling using consistent RMS-noise* (NUWS-UCR). Comparison to Equations (15) and (16) show that this weighting is distinct from both NUWS and apodization-weighting; the three techniques are

illustrated side-by-side in Figure S1. For an exponential sampling bias, $\sqrt{n(t_i)} \propto e^{-t_i/(2T_{SMP})}$, and a simple procedure to generate the NUS-UCR signal is to first calculate the average FID from

the mean of the $n(t_i)$ data points at each time point (Equation (15)) and then apply an exponential weighting function $e^{-t/(2T_{SMP})}$. Compared to NUWS, the additional linewidth introduced is halved.

Monte Carlo simulations of NUWS-UCR shown in Figure 7a confirm the preservation of spectral knowledge for amplitude in the frequency domain compared to NUS in the time domain; representative spectra generated during the Monte Carlo analysis are also shown for comparison. Figure S5 shows that the spectral knowledge for peak position and linewidth also fully track the intrinsic spectral knowledge of the time-domain NUS signal. This demonstrates a procedure that both theoretically and numerically preserves the knowledge in both domains for a non-uniformly sampled data set. Because the construction of the NUWS-UCR FID and the FT are linear, the NUS knowledge enhancement is also preserved when the signal is constructed from an arbitrary number of resonances. Figure S6 presents the results of Monte Carlo simulations analogous to those for the multiple resonances considered above (Figure 6) that demonstrate this.

3.3.2. Off-Grid Methods

3.3.2.1 Block Fourier Transform

NUWS still cleaves to the Nyquist grid. The more general approach is to dispense with the Nyquist grid altogether. The block Fourier Transform is one method to process data acquired off grid. Here we apply the BFT under conditions in which the data are relatively dense, with the sampling rate of the final points never falling below half that of the uniformly sampled data above.

Monte Carlo simulations of the NUS knowledge enhancement for spectra generated by the BFT applied to data acquired using off-grid, matched exponential NUS are shown in Figure 7b. These data were simulated with the same total number of data points (independent measures) as the NUWS example above. For comparison, the NUS FID was again fit directly in the time domain (NUS) and a reference US FID with the same number of measured time points was generated (US) and fit so that the NUS knowledge enhancement could be calculated.

There are two important features to note in Figure 7b. First, as one would intuit, the off-grid NUS in the time domain gives essentially identical NUS knowledge enhancement to on-grid NUS (Figure 7a). On-grid and off-grid NUS perform, for all practical purposes, equivalently. Second, as with NUWS, the BFT underperforms relative to the direct fit of the NUS time-domain data: only 40% of the NUS knowledge enhancement is retained.

3.3.2.2. Block Fourier Transform using Consistent RMS Noise

On its own, the BFT does not preserve the intrinsic spectral knowledge. Consideration of the noise characteristics analogous to those above for the NUWS-UCR argues that the FID should first be weighted by the square-root of the sampling density – a procedure that generates consistent RMS noise in the frequency domain from each unit-length time-domain

line segment. We refer to this method as the *block Fourier transform using consistent RMS noise* (BFT-UCR). Monte Carlo simulations summarized in Figure 7b and Figure S7 confirm the preservation of spectral knowledge in the frequency domain using the BFT-UCR compared to the direct fit of the off-grid NUS data in the time domain. Again, representative spectra generated during the Monte Carlo procedure are shown for comparison. Figure S6 verifies that the NUS knowledge enhancement is maintained for the BFT-UCR in the case of multiple resonances as well.

The BFT-UCR demonstrates a second procedure that preserves the intrinsic spectral knowledge of a non-uniformly sampled data set upon transformation to the frequency domain. The BFT-UCR is applicable to off-grid NUS, and although it was demonstrated here on simulated data that were relatively dense, we find that this can be relaxed, particularly in the later stages of the FID. The BFT-UCR may therefore find application in situations distinct from NUWS-UCR.

3.4. Using Consistent RMS Noise and Increased Signal-to-Noise Ratio

Together, the BFT-UCR and NUWS-UCR point to a general rule for the acquisition and processing of NMR data: maximum knowledge of the extracted NMR parameters occurs when the data are properly constructed using consistent root-mean-square noise – that is, when the RMS noise is equal at each time-domain point of the FID or each frequency-domain basis used to construct the spectrum. Both the BFT-UCR and NUWS-UCR generate spectra with equivalent spectral knowledge to that of the corresponding time-domain NUS signals. They therefore contain the maximum achievable spectral knowledge.

While the spectral knowledge of the extracted NMR parameters is perhaps the most relevant metric for assessing the sensitivity of NMR data, the importance of the SNR for identifying peaks in the frequency domain cannot be denied. There is still some controversy regarding the size of the NUS enhancement with respect to the signal-to-noise ratio.^{11,27} We posit that since the frequency domains under NUWS-UCR and BFT-UCR fully retain the intrinsic spectral knowledge of the original (unmanipulated) time-domain NUS data and that they are generated via Fourier transformation, that their SNR remains a fundamental measure of the quality of the NMR data. As shown in Figure 7, for spectra generated using NUWS-UCR and BFT-UCR and matched exponential bias with $t_{\max} = 3T_2$, the SNR is increased by 33% compared to the equivalent-time US spectra. However, the linewidths for the NUS data are 50% larger than for US, which changes the maximum SNR that could be obtained for each data set by the application of the respective matched filter.¹ When matched filters are applied to the NUWS-UCR, BFT-UCR, and US data, the SNR advantage for matched NUS increases to 45%. The increase in SNR under NUWS-UCR and BFT-UCR is larger than the corresponding 20% improvement observed by Waudby *et al.*,¹¹ but smaller than the 66% gain expected from the iSNR of Palmer *et al.*²⁷ Regardless, from the standpoint of both spectral knowledge and SNR, there is no doubt about the efficacy of NUS.

4. Conclusions

Monte Carlo simulations demonstrate that there is considerable benefit – quantified in terms of the intrinsic spectral knowledge for the peak amplitude – to weighted NUS in the indirect dimension(s) of NMR experiments. For an acquisition time of $3T_2$, NUS knowledge enhancements of 45% to 75% are obtained for sampling biases of 1 and 2, respectively, corresponding to time savings of 2-3 for equivalent sensitivity with US. While there is a concomitant decrease of up to 10%-20% in the spectral knowledge of the peak position and linewidth under NUS, this trade-off should be weighed against the increased ability to identify peaks and quantify the peak amplitude. There is, however, no guarantee that NUS will always outperform US when the goal is to assign intensity to individual peaks; in the case of partially resolved lines the intrinsic spectral knowledge falls below US for aggressive NUS biases (≥ 2). A reasonable compromise in most cases will be the use of matched NUS with data collected out to 2-3 T_2 , in which case the NUS knowledge enhancement is on the order of 45% for well-resolved resonances and still comparable to US for partially resolved lines. The increased spectral knowledge is not particularly sensitive to the precise bias or $t_{\max}$, so reasonable estimates are sufficient from a practical standpoint.

Two techniques for rigorously maintaining the full spectral knowledge when going from the time domain to the frequency domain have been introduced: non-uniform weighted sampling using consistent RMS noise (NUWS-UCR) and the Block FT using consistent RMS noise (BFT-UCR). These two techniques are applicable to on-grid and off-grid sampling, respectively. NUWS-UCR and the BFT-UCR illustrate how maximum knowledge of the NMR parameters is realized when the signals are constructed so that there is equivalent root-mean-square noise. Spectra generated with NUWS-UCR and BFT-UCR fundamentally contain the maximum achievable spectral knowledge.

NUWS-UCR and BFT-UCR are restricted to relatively densely sampled data and leave open the larger question of the most robust means to accomplish spectral reconstruction from sparsely sampled data. This is an area of considerable active research, and the Monte Carlo techniques used here should be of substantial use in quantifying the spectral knowledge in both the time and frequency domains for these cases as well. We recommend that Monte Carlo simulations be taken as the gold standard in assessing the quality of any proposed acquisition scheme and the fidelity of spectral reconstruction.

Supporting Information

Construction of time-domain NUWS, apodization-weighted, and NUWS-UCR FID's; Monte Carlo convergence; additional plots of NUS knowledge enhancements for peak amplitude, offset frequency, and linewidth.

Acknowledgements

Research reported in this publication was supported by NSF Grant CHE1710671 and NIH Grant GM097569.

References

- (1) Ernst, R. R.; Bodenhausen, G.; Wokaun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*; Clarendon Press ; Oxford University Press: Oxford Oxfordshire, New York, 1987.
- (2) Abragam, A. *The Principles of Nuclear Magnetism*; Clarendon Press: Oxford,, 1961.
- (3) Maly, T.; Debelouchina, G. T.; Bajaj, V. S.; Hu, K. N.; Joo, C. G.; Mak-Jurkauskas, M. L.; Sirigiri, J. R.; van der Wel, P. C. A.; Herzfeld, J.; Temkin, R. J.; Griffin, R. G. Dynamic Nuclear Polarization at High Magnetic Fields. *J. Chem. Phys.* **2008**, *128*.
- (4) Bowers, C. R.; Weitekamp, D. P. Para-Hydrogen and Synthesis Allow Dramatically Enhanced Nuclear Alignment. *J. Am. Chem. Soc.* **1987**, *109*, 5541-5542.
- (5) Lee, J. H.; Okuno, Y.; Cavagnero, S. Sensitivity Enhancement in Solution NMR: Emerging Ideas and New Frontiers. *J. Magn. Reson.* **2014**, *241*, 18-31.
- (6) Gan, Z. H.; Hung, I.; Wang, X. L.; Paulino, J.; Wu, G.; Litvak, I. M.; Gor'kov, P. L.; Brey, W. W.; Lendi, P.; Schiano, J. L.; Bird, M. D.; Dixon, L. R.; Toth, J.; Boebinger, G. S.; Cross, T. A. NMR Spectroscopy up to 35.2 T Using a Series-Connected Hybrid Magnet. *J. Magn. Reson.* **2017**, *284*, 125-136.
- (7) Kovacs, H.; Moskau, D.; Spraul, M. Cryogenically Cooled Probes - a Leap in NMR Technology. *Prog. Nucl. Magn. Reson. Spectrosc.* **2005**, *46*, 131-155.
- (8) Barna, J. C. J.; Laue, E. D.; Mayger, M. R.; Skilling, J.; Worrall, S. J. P. Exponential Sampling, an Alternative Method for Sampling in Two-Dimensional NMR Experiments. *J. Magn. Reson.* **1987**, *73*, 69-77.
- (9) Kumar, A.; Brown, S. C.; Donlan, M. E.; Meier, B. U.; Jeffs, P. W. Optimization of Two-Dimensional NMR by Matched Accumulation. *J. Magn. Reson.* **1991**, *95*, 1-9.
- (10) Rovnyak, D.; Sarcone, M.; Jiang, Z. Sensitivity Enhancement for Maximally Resolved Two-Dimensional NMR by Nonuniform Sampling. *Magn. Reson. Chem.* **2011**, *49*, 483-91.
- (11) Waudby, C. A.; Christodoulou, J. An Analysis of NMR Sensitivity Enhancements Obtained Using Non-Uniform Weighted Sampling, and the Application to Protein NMR. *J. Magn. Reson.* **2012**, *219*, 46-52.
- (12) Simon, B.; Kostler, H. Improving the Sensitivity of Ft-NMR Spectroscopy by Apodization Weighted Sampling. *J. Biomol. NMR* **2019**, *73*, 155-165.
- (13) Rovnyak, D.; Frueh, D. P.; Sastry, M.; Sun, Z. Y.; Stern, A. S.; Hoch, J. C.; Wagner, G. Accelerated Acquisition of High Resolution Triple-Resonance Spectra Using Non-Uniform Sampling and Maximum Entropy Reconstruction. *J. Magn. Reson.* **2004**, *170*, 15-21.
- (14) Sun, Z. Y.; Frueh, D. P.; Selenko, P.; Hoch, J. C.; Wagner, G. Fast Assignment of ¹⁵N-Hsqc Peaks Using High-Resolution 3d Hncocanh Experiments with Non-Uniform Sampling. *J. Biomol. NMR* **2005**, *33*, 43.
- (15) Frueh, D. P.; Sun, Z. Y.; Vosburg, D. A.; Walsh, C. T.; Hoch, J. C.; Wagner, G. Non-Uniformly Sampled Double-Trosy Hncanh Experiments for NMR Sequential Assignments of Large Proteins. *J. Am. Chem. Soc.* **2006**, *128*, 5757-63.
- (16) Mobli, M.; Stern, A. S.; Bermel, W.; King, G. F.; Hoch, J. C. A Non-Uniformly Sampled 4d Hcc(Co)Nh-Tocsy Experiment Processed Using Maximum Entropy for Rapid Protein Sidechain Assignment. *J. Magn. Reson.* **2010**, *204*, 160-164.
- (17) Hyberts, S. G.; Arthanari, H.; Wagner, G. Applications of Non-Uniform Sampling and Processing. *Top. Curr. Chem.* **2012**, *316*, 125-48.

- (18) Nowakowski, M.; Saxena, S.; Stanek, J.; Zerko, S.; Kozminski, W. Applications of High Dimensionality Experiments to Biomolecular NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **2015**, *90-91*, 49-73.
- (19) Coggins, B. E.; Venters, R. A.; Zhou, P. Radial Sampling for Fast NMR: Concepts and Practices over Three Decades. *Prog. Nucl. Magn. Reson. Spectrosc.* **2010**, *57*, 381-419.
- (20) Kazimierczuk, K.; Stanek, J.; Zawadzka-Kazimierczuk, A.; Kozminski, W. Random Sampling in Multidimensional NMR Spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **2010**, *57*, 420-434.
- (21) Mobli, M.; Hoch, J. C. Nonuniform Sampling and Non-Fourier Signal Processing Methods in Multidimensional NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *83*, 21-41.
- (22) Hoch, J. C.; Maciejewski, M. W.; Mobli, M.; Schuyler, A. D.; Stern, A. S. Nonuniform Sampling and Maximum Entropy Reconstruction in Multidimensional NMR. *Acc. Chem. Res.* **2014**, *47*, 708-17.
- (23) Hyberts, S. G.; Arthanari, H.; Robson, S. A.; Wagner, G. Perspectives in Magnetic Resonance: NMR in the Post-Fft Era. *J. Magn. Reson.* **2014**, *241*, 60-73.
- (24) Kazimierczuk, K.; Orekhov, V. Non-Uniform Sampling: Post-Fourier Era of NMR Data Collection and Processing. *Magn. Reson. Chem.* **2015**, *53*, 921-6.
- (25) Hoch, J. C. Beyond Fourier. *J. Magn. Reson.* **2017**, *283*, 117-123.
- (26) Golowicz, D.; Kasprzak, P.; Orekhov, V.; Kazimierczuk, K. Fast Time-Resolved NMR with Non-Uniform Sampling. *Prog. Nucl. Magn. Reson. Spectrosc.* **2020**, *116*, 40-55.
- (27) Palmer, M. R.; Suiter, C. L.; Henry, G. E.; Rovnyak, J.; Hoch, J. C.; Polenova, T.; Rovnyak, D. Sensitivity of Nonuniform Sampling NMR. *J. Phys. Chem. B* **2015**, *119*, 6502-15.
- (28) Paramasivam, S.; Suiter, C. L.; Hou, G.; Sun, S.; Palmer, M.; Hoch, J. C.; Rovnyak, D.; Polenova, T. Enhanced Sensitivity by Nonuniform Sampling Enables Multidimensional MAS NMR Spectroscopy of Protein Assemblies. *J. Phys. Chem. B* **2012**, *116*, 7416-27.
- (29) Qiang, W. Signal Enhancement for the Sensitivity-Limited Solid State NMR Experiments Using a Continuous, Non-Uniform Acquisition Scheme. *J. Magn. Reson.* **2011**, *213*, 171-175.
- (30) Suiter, C. L.; Paramasivam, S.; Hou, G. J.; Sun, S. J.; Rice, D.; Hoch, J.; Rovnyak, D.; Polenova, T. Sensitivity Gains, Linearity, and Spectral Reproducibility in Nonuniformly Sampled Multidimensional MAS NMR Spectra of High Dynamic Range. *J. Biomol. NMR* **2014**, *59*, 57-73.
- (31) Porat, G.; Goldbourt, A. Assessment of Non-Uniform Sampling Schemes in Solid State NMR of Bacteriophage Viruses. *Isr. J. Chem.* **2019**, *59*, 1027-1038.
- (32) Bracewell, R. N. *The Fourier Transform and Its Applications*; McGraw-Hill: New York,, 1965.
- (33) Marshall, A. G.; Verdun, F. R. *Fourier Transforms in NMR, Optical, and Mass Spectrometry*; Elsevier: Amsterdam, 1990.
- (34) Donoho, D. L.; Johnstone, I. M.; Stern, A. S.; Hoch, J. C. Does the Maximum Entropy Method Improve Sensitivity? *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 5066.
- (35) Sibisi, S.; Skilling, J.; Brereton, R. G.; Laue, E. D.; Staunton, J. Maximum-Entropy Signal-Processing in Practical NMR-Spectroscopy. *Nature* **1984**, *311*, 446-447.
- (36) Hyberts, S. G.; Heffron, G. J.; Tarragona, N. G.; Solanky, K.; Edmonds, K. A.; Luithardt, H.; Fejzo, J.; Chorev, M.; Aktas, H.; Colson, K.; Falchuk, K. H.; Halperin, J. A.; Wagner, G. Ultrahigh-Resolution H-1-C-13 Hsqc Spectra of Metabolite Mixtures Using

- Nonlinear Sampling and Forward Maximum Entropy Reconstruction. *J. Am. Chem. Soc.* **2007**, *129*, 5108-5116.
- (37) Hyberts, S. G.; Milbradt, A. G.; Wagner, A. B.; Arthanari, H.; Wagner, G. Application of Iterative Soft Thresholding for Fast Reconstruction of NMR Data Non-Uniformly Sampled with Multidimensional Poisson Gap Scheduling. *J. Biomol. NMR* **2012**, *52*, 315-27.
- (38) Kazimierczuk, K.; Orekhov, V. Y. Accelerated NMR Spectroscopy by Using Compressed Sensing. *Angew. Chem. Int. Ed.* **2011**, *50*, 5556-5559.
- (39) Ying, J. F.; Delaglio, F.; Torchia, D. A.; Bax, A. Sparse Multidimensional Iterative Lineshape-Enhanced (Smile) Reconstruction of Both Non-Uniformly Sampled and Conventional NMR Data. *J. Biomol. NMR* **2017**, *68*, 101-118.
- (40) Sun, S. J.; Gill, M.; Li, Y. F.; Huang, M.; Byrd, R. A. Efficient and Generalized Processing of Multidimensional Nus NMR Data: The NESTA Algorithm and Comparison of Regularization Terms. *J. Biomol. NMR* **2015**, *62*, 105-117.
- (41) Coggins, B. E.; Zhou, P. High Resolution 4-D Spectroscopy with Sparse Concentric Shell Sampling and Fft-Clean. *J. Biomol. NMR* **2008**, *42*, 225-239.
- (42) Coggins, B. E.; Werner-Allen, J. W.; Yan, A.; Zhou, P. Rapid Protein Global Fold Determination Using Ultrasparse Sampling, High-Dynamic Range Artifact Suppression, and Time-Shared Noesy. *J. Am. Chem. Soc.* **2012**, *134*, 18619-18630.
- (43) Blum, R. L.; Rovny, J.; Loria, J. P.; Barrett, S. E. Reaching the Sparse-Sampling Limit for Reconstructing a Single Peak in a 2d NMR Spectrum Using Iterated Maps. *J. Biomol. NMR* **2019**, *73*, 545-560.
- (44) Orekhov, V. Y.; Jaravine, V. A. Analysis of Non-Uniformly Sampled Spectra with Multi-Dimensional Decomposition. *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *59*, 271-292.
- (45) Hansen, D. F. Using Deep Neural Networks to Reconstruct Non-Uniformly Sampled NMR Spectra. *J. Biomol. NMR* **2019**, *73*, 577-585.
- (46) Yoon, J. W.; Godsill, S.; Kupce, E.; Freeman, R. Deterministic and Statistical Methods for Reconstructing Multidimensional NMR Spectra. *Magn. Reson. Chem.* **2006**, *44*, 197-209.
- (47) Palmer, M. R.; Wenrich, B. R.; Stahlfeld, P.; Rovnyak, D. Performance Tuning Non-Uniform Sampling for Sensitivity Enhancement of Signal-Limited Biological NMR. *J. Biomol. NMR* **2014**, *58*, 303-314.
- (48) Press, W. H.; Teukolsky, S. A.; Vetterling, W.; Flannery, B. P. *Numerical Recipes in C : The Art of Scientific Computing*; 2 ed.; Cambridge University Press: Cambridge Cambridgeshire ; New York, 1992.
- (49) Ernst, R. R. Sensitivity Enhancement in Magnetic Resonance. *Adv. Magn. Reson.* **1966**, *2*, 1.
- (50) Matson, G. B. Signal Integration and Signal-to-Noise Ratio in Pulsed NMR Relaxation Measurements. *J. Magn. Reson.* **1977**, *25*, 477-480.
- (51) Williams, B. W.; Rovnyak, D. A Closed-Form Solution for the Optimal Time Evolution of an Exponentially Decaying Signal with Thermal Noise. *Concepts Magn. Reson. A* **2014**, *43*, 54-56.
- (52) Dudley, J. A.; Park, S.; MacDonald, M. E.; Fetene, E.; Smith, C. A. Resolving Overlapped Signals with Automated FitNMR Analytical Peak Modeling. *J. Magn. Reson.* **2020**, *in press*.

- (53) Hyberts, S. G.; Robson, S. A.; Wagner, G. Exploring Signal-to-Noise Ratio and Sensitivity in Non-Uniformly Sampled Multi-Dimensional NMR Spectra. *J. Biomol. NMR* **2013**, *55*, 167-78.
- (54) Zambrello, M. A.; Craft, D. L.; Hoch, J. C.; Rovnyak, D.; Schuyler, A. D. The Influence of the Probability Density Function on Spectral Quality in Nonuniformly Sampled Multidimensional NMR. *J. Magn. Reson.* **2020**, *311*.
- (55) Jones, J. A.; Hodgkinson, P.; Barker, A. L.; Hore, P. J. Optimal Sampling Strategies for the Measurement of Spin-Spin Relaxation Times. *J. Magn. Reson., Ser B* **1996**, *113*, 25-34.
- (56) Hoch, J. C.; Maciejewski, M. W.; Filipovic, B. Randomization Improves Sparse Sampling in Multidimensional NMR. *J. Magn. Reson.* **2008**, *193*, 317-20.
- (57) Kupce, E.; Freeman, R. Fast Multidimensional NMR: Radial Sampling of Evolution Space. *J. Magn. Reson.* **2005**, *173*, 317-321.
- (58) Mobli, M.; Stern, A. S.; Hoch, J. C. Spectral Reconstruction Methods in Fast NMR: Reduced Dimensionality, Random Sampling and Maximum Entropy. *J. Magn. Reson.* **2006**, *182*, 96-105.
- (59) Kazimierczuk, K.; Zawadzka, A.; Kozminski, W. Optimization of Random Time Domain Sampling in Multidimensional NMR. *J. Magn. Reson.* **2008**, *192*, 123-130.
- (60) Hyberts, S. G.; Takeuchi, K.; Wagner, G. Poisson-Gap Sampling and Forward Maximum Entropy Reconstruction for Enhancing the Resolution and Sensitivity of Protein NMR Data. *J. Am. Chem. Soc.* **2010**, *132*, 2145-7.
- (61) Mobli, M.; Maciejewski, M. W.; Schuyler, A. D.; Stern, A. S.; Hoch, J. C. Sparse Sampling Methods in Multidimensional NMR. *PCCP* **2012**, *14*, 10835-10843.
- (62) Lin, E. C.; Opella, S. J. Sampling Scheme and Compressed Sensing Applied to Solid-State NMR Spectroscopy. *J. Magn. Reson.* **2013**, *237*, 40-48.
- (63) Aoto, P. C.; Fenwick, R. B.; Kroon, G. J. A.; Wright, P. E. Accurate Scoring of Non-Uniform Sampling Schemes for Quantitative NMR. *J. Magn. Reson.* **2014**, *246*, 31-35.
- (64) Delaglio, F.; Walker, G. S.; Farley, K. A.; Sharma, R.; Hoch, J. C.; Arbogast, L. W.; Brinson, R. G.; Marino, J. P. Non-Uniform Sampling for All: More NMR Spectral Quality, Less Measurement Time. *Am. Pharm. Rev.* **2017**, *20*, 339681.
- (65) Maciejewski, M. W.; Schuyler, A. D.; Hoch, J. C. Practical Nonuniform Sampling and Non-Fourier Spectral Reconstruction for Multidimensional NMR. *Methods Mol. Biol.* **2018**, *1688*, 341-352.
- (66) *Mathematica*, version 11.2; software for technical computation; Wolfram Research, Inc.: Champaign, Illinois, 2017.

