Solid-State NMR Spectroscopy

Rapid 3D MAS NMR Spectroscopy at Critical Sensitivity**

Yoh Matsuki, Matthew T. Eddy, Robert G. Griffin, and Judith Herzfeld*

Abstract

In solution NMR, more efficient acquisition has relied on non-uniform sampling (NUS), which has been successfully applied to multidimensional experiments on larger systems. Although extending NUS to MAS NMR would be of enormous practical importance, the application of conventional NUS methods to MAS NMR has been limited by the specific problem of accurately modeling weak signals in noisy spectra in addition to the general problems of quantitative spectral reconstruction and slow computation.

Herein, we address these challenges with SIFT (spectroscopy integration from time-domain artifacts) for NUS MAS NMR, recorded a 3D NCOCX spectrum (Figure 1a) of a microcrystalline, uniformly $^{15}$N,$^{13}$C-labeled sample of the beta domain of protein G (GB1) at high digital resolution (1.1 ppm for F1, 0.7 ppm for F2, before zero-filling). For both t1 and t2, the dwell time was synchronized to three times the rotor period, $3/v_p$ (bandwidth equal to $v_p/3$ Hz), in order to fold the a priori knowledge of “dark” regions in the frequency domain; that is, those regions known to contain no NMR signals. The frequency domain information, assimilated by a very rapid computational process, obviates some time-domain sampling with no sacrifice in resolution and no modeling bias.

We previously used SIFT to process 2D NUS $^{15}$N-HSQC solution data, where dark regions created by over-sampling were utilized to replace up to 25% of the uniform time-domain data points. We demonstrate the effectiveness of the SIFT method in solids, using dark regions resulting from the need for rotor-synchronized sampling in the indirect dimensions. Unlike other NUS data processing methods that actively model signals to reconstruct a spectrum, SIFT suppresses the sampling noise by using only definitive information from the dark spectral areas. Thus, SIFT avoids bias from subjective discrimination between weak signals and noise, and reconstructs missing time data points with high fidelity, as if they had been actually recorded. These favorable properties make SIFT uniquely suited for processing NUS data in the sensitivity-limited regime.

To demonstrate the application of SIFT to NUS MAS NMR, we recorded a 3D NCOCX spectrum (Figure 1a) of a microcrystalline, uniformly $^{15}$N,$^{13}$C-labeled sample of the beta domain of protein G (GB1) at high digital resolution (1.1 ppm for F1, 0.7 ppm for F2, before zero-filling). For both t1 and t2, the dwell time was synchronized to three times the rotor period, $3/v_p$ (bandwidth equal to $v_p/3$ Hz), in order to fold the a priori knowledge of “dark” regions in the frequency domain; that is, those regions known to contain no NMR signals. The frequency domain information, assimilated by a very rapid computational process, obviates some time-domain sampling with no sacrifice in resolution and no modeling bias.

We previously used SIFT to process 2D NUS $^{15}$N-HSQC solution data, where dark regions created by over-sampling were utilized to replace up to 25% of the uniform time-domain data points. We demonstrate the effectiveness of the SIFT method in solids, using dark regions resulting from the need for rotor-synchronized sampling in the indirect dimensions. Unlike other NUS data processing methods that actively model signals to reconstruct a spectrum, SIFT suppresses the sampling noise by using only definitive information from the dark spectral areas. Thus, SIFT avoids bias from subjective discrimination between weak signals and noise, and reconstructs missing time data points with high fidelity, as if they had been actually recorded. These favorable properties make SIFT uniquely suited for processing NUS data in the sensitivity-limited regime.

To demonstrate the application of SIFT to NUS MAS NMR, we recorded a 3D NCOCX spectrum (Figure 1a) of a microcrystalline, uniformly $^{15}$N,$^{13}$C-labeled sample of the beta domain of protein G (GB1) at high digital resolution (1.1 ppm for F1, 0.7 ppm for F2, before zero-filling). For both t1 and t2, the dwell time was synchronized to three times the rotor period, $3/v_p$ (bandwidth equal to $v_p/3$ Hz), in order to fold the a priori knowledge of “dark” regions in the frequency domain; that is, those regions known to contain no NMR signals. The frequency domain information, assimilated by a very rapid computational process, obviates some time-domain sampling with no sacrifice in resolution and no modeling bias.

Figure 1. a) The projection down the direct acquisition axis (F3) of the NCOCX spectrum of GB1. The full spectral range is shown for both the $^{15}$N (F1) and $^{13}$C (F2) dimensions. b) The dots represent the NUS schedule comprising 608 quasi-random on-grid points with exponentially decaying probability density. The shaded area corresponds to a truncated set of 608 (32 x 19) uniform points. A full dataset with 2048 (64 x 32) samples was acquired as a reference and time samples were omitted according to the schedule to form the NUS dataset. Forty scans were averaged per t1/t2 sample. The NUS schedule was generated by the sampling scheduler available at http://sbtools.ucdavis.edu/nmr/sample_scheduler/. The dwell time was 270 μs for t1 and t2 at a MAS frequency of 11.1 kHz.
spinning frequency \( v_s \) was chosen to avoid rotational resonance owing to overlap of the sidebands of carbonyl \(^{13}\text{C} \) signals with the aromatic and aliphatic signals in the acquisition dimension (F3). Because of these constraints, the bandwidth in the indirect acquisition dimensions (F1, F2) left spectral regions known to be devoid of signals. Whereas those “dark” regions are conventionally neglected, SIFT actively uses them.

To model critical sensitivity, we intentionally underpacked the sample rotor such that only about 50% of the active volume was filled. With a molecular weight of about 6.4 kDa for GB1, the NMR sensitivity for this half-packed sample would correspond to that of a fully packed circa 13 kDa microcrystalline protein at the same density. Signal averaging took about 7 min per 11/2 sample for an acceptable S/N ratio, whilst extended signal evolution was also required for suitable linewidths.

The NUS schedule employed (Figure 1b) omits more than 70% of the time domain samples in the full, uniform grid of 64(11) x 32(12) time points. The input for SIFT processing comprises the NUS schedule, the acquired time-domain data, and specification of the dark spectral regions. The processing, which involves no user intervention or parameter tuning, took about 2 min on a single processor. The SIFTed time domain data may be transformed and phased as though directly acquired.

Figure 2 shows that the S/N degradation owing to NUS (d) is almost perfectly reversed after SIFT-processing (c), leading to S/N that is nearly the same as that of the fully acquired spectrum with more than three times as many acquired points (a). This effect is significantly better than the square root of the number of NUS data (Figure 2d) resulted in 18 (29%) of the 62 weaker peaks being lost (that is, intensity below the threshold of 3 \( \sigma \)). The correlation coefficient of the classical discrete Fourier transform (DFT) of the NUS data to the original data was 0.995 overall and 0.877 for the “weak” signals. The minimum contour line is at 1\% of the tallest peak in each panel. 1D traces at dashed lines are also shown, for which the vertical scales are the same for all panels. The slice shown contains signals with about 1–2 times the median peak intensity and a S/N ratio of between 8.4 and 15.6. Slices with weaker signals are shown in the Supporting Information, Figure S1.

In addition to excellent S/N, SIFT offers high fidelity. Line shape distortion of the type seen in MaxEnt-processed spectra is not observed in the SIFTed spectrum, despite the much lower intrinsic S/N of the present data. We also found that SIFT yields more accurate peak intensities and positions than those obtained by linear prediction (LP) of a truncated uniformly sampled (US) dataset. For example, Figure 2c shows that LP does not accurately restore crowded spectral regions even with a reasonable number of coefficients (8). This effect is especially clear in the F1 slice. Further instances are illustrated by the additional F1 slices in Figure 3. In all the spectral regions where LP failed, the SIFT spectrum reproduced the full data very well. Overall, SIFT rendered peak frequencies with RMS errors of 0.13, 0.11, and 0.09 ppm in F1, F2, and F3, respectively (see the Supporting Information, Figure S2).

To assess the accuracies of peak intensities, the intensities in the SIFTed spectrum and the linear predicted spectrum versus the “true” peak intensities obtained from the fully acquired dataset were plotted (Figure 4). The excellent linearity obtained with SIFT (Figure 4a–d) demonstrates accurate relative peak intensities. The correlation coefficient was 0.995 overall and 0.877 for the “weak” signals. The

---

These are not the final page numbers!
The vertical scale is the same throughout each column. The asterisks mark the unresolved peaks in the LP spectra.

**Figure 3.** Representative F1 slices through crowded spectral regions. Each column compares the spectrum of the full data (top), LP of truncated US data (middle), and SIFT-processed NUS data (bottom). From the left to right, slices were taken at $\delta = 174.7, 176.1$, and 179.8 ppm.

**Figure 4.** Signal intensities observed in the spectrum of a) FT-processed NUS data and b) processed US data versus fully sampled reference data. The linear regression shown with a solid line is slightly different between (a) and (b). Dashed lines flanking the regression line show the intrinsic noise width in the reference spectrum. Dashed lines along the x and y axes show the floor of noise standard deviation. The medians of all the observed signal intensities in the reference, SIFT-processed, and linear-predicted spectra (marked by arrows) were about 110, 100, and 95, respectively.

The dynamic range was about 13 in this example, but the superb linearity of SIFT signal intensities over a dynamic range of about 100 has been shown previously.[24] The variation of the observed peak intensities was mostly within the intrinsic noise-width of the dataset (ca. ±25), which is indicated by the dashed lines flanking the regression line. This observation indicates the absence of intensity bias in SIFT processing. The accuracy remained high for the weakest signals, thus confirming the robustness of the SIFT process in the presence of formidable noise. In contrast, LP tends to reproduce large signals with less accuracy (Figure 4b).

In conclusion, by using the noise-tolerant SIFT process,[24] we have extended the applicability of high-dimensional NUS-NMR methodology to data with the marginal sensitivity that is typical of MAS NMR of biological macromolecules. Quick SIFT processing (ca. 2 min) of NUS data yielded a high-quality 3D spectrum without any calibration or parameter optimization. After SIFT processing, the reduced number of time samples in NUS did not appreciably decrease the S/N relative to that for uniformly acquired reference data. Meanwhile, the measurement time was reduced by a factor of about 3.4. These results suggest that a 3D NOCCX-type MAS experiment can be recorded at sufficient sensitivity to resolve single nuclear sites for circa 25 kDa proteins in a reasonable period of time. The approach demonstrated herein requires no special hardware and will expedite experiments similarly on all FT spectrometers. For example, if a tenfold sensitivity gain were available by dynamic nuclear polarization (DNP),[25] the above high-resolution 3D experiment would be possible for 250 kDa proteins. Moreover, the exquisite accuracy of SIFT signal frequencies and intensities paves the way for quantitative structural and dynamical investigations in noisy systems that have frustrated all other reported NUS processing methods. Thus, SIFT will significantly expand opportunities for high-dimensional MAS NMR experiments in studies of large molecules and molecular assemblies of biological and medical importance.

**Experimental Section**

Uniformly $^{13}$C and $^{15}$N labeled GB1 was prepared according to previously published procedures[26,27] as described in the Supporting Information. The NMR experiment was performed on a custom-built 500 MHz (1H frequency) spectrometer equipped with a solenoid-coil 3.2 mm MAS system (Revolution NMR, Fort Collins, CO). The sampling schedule is converted into a text-based list that is read by the pulse program and control macro to set respective delays (courtesy of Dr. P. van der Wel, University of Pittsburgh). Details on the NMR parameters are given in the Supporting Information.

MATLAB scripts for SIFT processing are available at http://www.brandeis.edu/~herzfeld/SIFT. The signal-containing “bright” region was $\delta = 108.2-152.2$ ppm ($^{13}$C, $^{15}$N) and $\delta = 169.3-183.9$ ppm ($^{1}H$, $^{12}$C), which was established from observing experiments for the corresponding nuclei. The number of SIFT cycles was 10, which took about 2 min. Further details on processing parameters and spectral analysis are given in the Supporting Information.

Received: June 1, 2010
Published online: \_______, 2010

**Keywords:** data acquisition methods · non-uniform sampling · protein structures · NMR spectroscopy

---

Communications


These are not the final page numbers!
Communications

Solid-State NMR Spectroscopy

Y. Matsuki, M. T. Eddy, R. G. Griffin, J. Herzfeld*

Rapid 3D MAS NMR Spectroscopy at Critical Sensitivity

Sensitive SIFTing: Multidimensional non-uniform sampling (NUS) NMR spectroscopy is extended to the severely sensitivity-limited regime typical of MAS NMR of biomacromolecules by the use of spectroscopy by the integration of frequency and time domain information (SIFT). A 3D MAS NMR experiment is expedited (here three-to-fourfold) without losing sensitivity or resolution.

Festkörper-NMR-Spektroskopie

Y. Matsuki, M. T. Eddy, R. G. Griffin, J. Herzfeld*

Rapid 3D MAS NMR Spectroscopy at Critical Sensitivity

Mehrdimensionale NMR-Spektroskopie mit stichprobenhafter Datenaufnahme (non-uniform sampling, NUS) eignet sich auch für die wenig empfindliche MAS-NMR-Spektroskopie von Biomakromolekülen. Durch Spektroskopie unter Integration der Information aus Frequenz- und Zeitdomäne (SIFT) können 3D-MAS-NMR-Experimente ohne Verlust an Empfindlichkeit und Auflösung deutlich schneller ausgeführt werden (siehe Beispiel).