Published in final edited form as:

J Magn Reson. 2017 October ; 283: 117–123. doi:10.1016/j.jmr.2017.03.017.

Beyond Fourier

Jeffrey C. Hoch

Department of Molecular Biology and Biophysics, UConn Health, 263 Farmington Ave.,
Farmington, CT 06030-3305 USA

Abstract

Non-Fourier methods of spectrum analysis are gaining traction in NMR spectroscopy, driven by their utility for processing nonuniformly sampled data. These methods afford new opportunities for optimizing experiment time, resolution, and sensitivity of multidimensional NMR experiments, but they also pose significant challenges not encountered with the discrete Fourier transform. A brief history of non-Fourier methods in NMR serves to place different approaches in context. Non-Fourier methods reflect broader trends in the growing importance of computation in NMR, and offer insights for future software development.

Graphical abstract

Weakness is a hindrance, and a virtue

As working NMR spectroscopists, we risk becoming inured to just how remarkable NMR is. NMR spectroscopy is arguably the most versatile method yet devised for probing matter, able to probe solids, liquids and gasses. It is remarkably non-invasive, and can be used with homogenous or heterogeneous systems. It interacts only weakly with systems being analyzed, and in the words of the Hoosier engineer Ed Purcell the interaction with biological systems (for example in MRI) is “so small you can put it in your eye” (personal communication). This light touch makes NMR ideal for reporting on weakly interacting or marginally stable systems, for example the dynamically disordered yet functional proteins that make up the “Dark Proteome” [1, 2]. NMR is even more remarkable in the rare circumstances where it can be said to perturb the system being investigated beyond the manifold of nuclear spin states, for example the nuclear spin-state dependent radical

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.

recombination reactions that lead to the CIDNP effect [3], or in magnetic cooling at sub-milliKelvin temperatures [4].

An especially remarkable property of NMR is its range. It reports on structural details at length scales from that of oil well fields[5] to a small fraction of the length of a hydrogen bond[6, 7]. It is sensitive to dynamics on time scales spanning picoseconds to hours. The weak interaction of static magnetic and RF fields with matter enable NMR at both ends of the time and length scales, for example noninvasively penetrating the human body in MRI, detecting a fractional Ångström change in an internuclear distance, or sensing the nanosecond time-scale reorientation of a nuclear spin pair.

This weak interaction between analyst and analyte that is a defining characteristic of NMR is also its Achilles' Heel. Though NMR yields exquisite detail, it remains an exquisitely insensitive technique. In contrast to fluorescence methods able to detect single molecules, commercial NMR instruments typically require 10^{12} nuclear spins to yield a detectable signal (heroic achievements in detecting a single nuclear spin notwithstanding[8]). The most challenging problems for NMR test the very limits of sensitivity, and the quest for improved sensitivity is the driving force behind technical marvels such as ultra-high field superconducting magnets and cryogenically-cooled RF probes. We seek sensitivity gains in every facet of the NMR experiment.

The Spectrometer Gains a Computer

The quest for sensitivity doesn't end with the magnet or the probe. Since the simple yet profound proposal by Ernst and Anderson[9] that the NMR spectrum can easily be determined by recording the time-varying response of the spins to a perturbation, an equally important aspect of the NMR experiment is "spectrum analysis", the mathematical process of analyzing the time series recorded by the detector to identify and characterize the constituent frequency components. For Ernst and Anderson, spectrum analysis meant the discrete Fourier transform (DFT). Remarkably, they were unaware in 1966 of the Cooley-Tukey fast Fourier transform (FFT) algorithm [10], which was published the prior year. Given that their procedure required transferring data from the spectrometer via punched paper tape, to punched cards, then physically carrying the data to a nearby IBM service center to compute the spectrum, their proposal that Fourier transform NMR (FT-NMR) would accelerate NMR spectroscopy was fairly audacious. An interesting historical note is that Gauss was aware of the symmetries that form the basis for the FFT algorithm. Evidence is found in Gauss's notebooks, which can be viewed at the Göttingen Observatory on Geismar Landstraße in Göttingen. As a contemporary of Fourier's, it seems unlikely that Gauss considered the sums of discrete trigonometric functions to be "Fourier series".

Limitations of the DFT were apparent from the very beginning of FT-NMR. Most significant are the challenges obtaining high resolution spectra from short data records and the tradeoff between sensitivity and resolution. Ernst famously introduced the concept of matched filter apodization to the NMR community, though the concept was published at least as far back as 1946 [11], and there is evidence that Norbert Wiener [12] (in 1930) and Albert Einstein [13] (in 1913) had arrived at the result independently. Matched-filter apodization results in

optimal signal-to-noise (SNR) ratio but at the expense of diminished resolution through concomitant line broadening. In the early 1980's, linear prediction (LP) models, which were being widely used in parametric spectrum analysis, "organically" materialized in commercial NMR software – FT-NMR from Hare Research and NMR1 and NMR2 from New Methods Research, and the embedded software provided with commercial NMR spectrometers – in the form of LP extrapolation of the signal beyond the measured interval. The behind-the-scenes contributions to these developments from Andy Byrd and Marc-André Delsuc remain undocumented, but were in many respects seminal. LP extrapolation affords additional flexibility in the design of apodization functions applied prior to the DFT, and is especially useful for avoiding truncation artifacts associated with short data records that are padded with zeros (zero-filled) to increase the digital resolution of the DFT spectrum. Although LP extrapolation remains a widely-used tool for extending the indirect dimensions of multidimensional experiments, because it implicitly models the data as a sum of exponentially decaying sinusoids, it is subject to false positives, and bias when the signal decay is not exponential, or the noise is not random [14].

Beyond Fourier

Around the same time that LP extrapolation was gaining acceptance, a group led by Ernest Laue and John Skilling at Cambridge University applied a nonparametric method of spectrum analysis for analyzing NMR data [15] that had previously found success in analyzing geophysical and astronomical data. This "maximum entropy method" evolved from the information-theoretic concept of entropy as a measure of missing information introduced by Claude Shannon [16], with important subsequent contributions from John P. Burg [17] and Edwin T. Jaynes [18]. In the inverse approach to spectrum analysis, entropy is used as a regularization functional while ensuring that the inverse DFT of the reconstructed spectrum is consistent with the empirical data. This is in principle solvable using standard numerical approaches, though fixed-point and steepest-descent methods proved flawed. That changed with the development of the "Cambridge Algorithm" by Skilling and Bryan [19], which was the first robust and efficient algorithm capable of computing "MaxEnt" reconstructions of NMR spectra containing 10^5 or more unknowns. Extensions of the Shannon entropy to a form suitable for hypercomplex data [20, 21] encountered in NMR and algorithmic refinements resulted in methods well-suited for modern multidimensional NMR spectroscopy.

MaxEnt reconstruction was not widely adopted, no doubt because of a dearth of software and because it is less intuitive than LP extrapolation. It's also likely that the claims that MaxEnt is capable of simultaneously improving sensitivity and resolution were deemed by many to be dubious – as they remain today. However the facile ability of MaxEnt to obtain high-resolution spectra from nonuniformly sampled data ultimately provided the incentive for wider adoption. But the power of nonuniform sampling, though it was demonstrated using MaxEnt in a pair of seminal papers from Laue and colleagues [22, 23], was not fully appreciated by the NMR community until Ray Freeman and Eriks Kupče [24, 25] introduced an approach based on sampling along radial vectors in the indirect time dimensions and back-projection for computing the spectra. Based on concepts employed in computed tomography for reconstructing three-dimensional objects from line projections –

inverting the Radon transform – the dramatic time savings for three dimensional NMR experiments and improved resolution were widely embraced, and spun off methods such as APSY [26] and HIFI [27] that use radial sampling but eschew reconstruction of the multidimensional spectrum by performing analysis in the projected dimension(s). The connection between these approaches and the nonuniform sampling (NUS) approach employed by Laue and colleagues remained underappreciated until the demonstration that MaxEnt reconstruction could be used to invert the Radon transform for radially-sampled data [28] and that the radial FT is equivalent to weighted back-projection [29]. Thus the projection reconstruction (PR) approach was shown to be mainly a specific sampling strategy, and the artifacts characteristic of spectra computed from radially-sampled data are characteristic of the sampling scheme, and not the result of specific reconstruction techniques. Developing and characterizing optimal sampling schemes that exhibit high efficiency, sensitivity, and produce few artifacts became and remains an active area of research [30–32].

Outside the field of NMR, parallel developments in nonuniform sampling (a.k.a. sparse recovery) were taking place. An early germ of an idea underlying sparse recovery was planted by B. F. “Tex” Logan, who showed in his Ph.D. thesis [33] (completed as a night student at Columbia University while working at Bell Labs; his other night job was as a celebrated bluegrass musician, Figure 1) that under certain conditions it is possible to *perfectly* recover a signal corrupted by noise or jamming, by minimizing the l_1 -norm of the spectrum. David Donoho [34] deserves credit for recognizing the full implications of Logan’s Theorem for recovering signals when measurements are *missing*, not merely corrupted. Donoho [35], his student Emanuel Candes, and Terrance Tao [36] embraced the term “compressed sensing” (CS) to describe the use of l_1 -norm minimization to recover a signal from fewer samples than are stipulated by the Nyquist Theorem. A fundamental tenet of CS for spectrum analysis is that sparse recovery is possible when the spectrum to be recovered is sparse, or “nearly black”: only a small fraction of the spectrum contains signals [37]. Only some of the complexities of CS are directly relevant for NMR spectroscopy, where the data and “image” domains are related by Fourier transformation, and theory needed to apply CS to NMR applications remains an active area of exploration [38, 39].

Theoreticians are progressively tightening the minimum bound on the number of samples needed to recover a spectrum, which depends on the sparsity of the spectrum to be recovered and characteristics of the sampling scheme employed. Donoho and Tanner [39] showed that the probability of accurate recovery exhibits a phase transition as the number of samples increases (the sampling “coverage”; Figure 2). The precise location of the phase transition depends on the sampling scheme, and the functional used to regularize the recovered spectrum, but the existence of a phase transition is a universal property of any “reasonable” sampling scheme that is sufficiently incoherent. The existence of these phase transitions has potentially profound implications for sensitivity of NMR experiments, because it implies that for some experiments small increases in sparsity (e.g. through higher magnetic field) could traverse the phase transition, from the “failure” to “success” regions of the recovery probability. For such problems the increase in field strength could result in a gain of sensitivity far exceeding the gain expected from the Curie Law increase in magnetization.

It bears emphasizing that the field of CS didn't begin in the early 2000's with the coining of the term "compressed sensing". There is a decades-long history of the use of sparse sampling in NMR, starting with the seminal contributions of Laue and colleagues in the early 1980's, and an even longer history of the use of a variety of ingenious approaches to dealing with missing data. For example, the "Burg maximum entropy method"[17] was inspired partly by the nasty tendency of the pens in field seismographs to run out of ink, leaving gaps in the record. CS also joins with another long-running thread, in the realm of image processing, where iterative thresholding algorithms have a long history. It turns out that iterative soft thresholding minimizes the l_1 -norm [40]. An iterative thresholding algorithm developed for image processing in 1931 by van Cittert [41] is a direct antecedent to iterative soft thresholding algorithms used today to perform CS.

There remains considerable room for theoretical developments in CS, as it applies to NMR. For example it was only recently realized by Monajemi that the hypercomplex algebra so familiar to NMR spectroscopists demands generalization of notions of coherence beyond the real-valued formulations employed in most theories of CS [42]. Similarly, the implications of the *uniform* sampling employed in the direct (acquisition) dimension of multidimensional NMR experiments for sparse recovery were also discovered by Monajemi, providing a theoretical explanation for the empirical observation of the challenges of employing NUS in a single indirect dimension [14]. Further theoretical developments can be expected to guide improvements in experiment design for multidimensional NMR.

MaxEnt, CS, and BPR belong to the class of nonparametric methods of spectrum analysis, in that they make no assumptions about the nature of the signals. No discussion of non-Fourier methods in NMR would be complete, however, without mentioning the much broader range of non-Fourier methods that have been successfully applied to NMR; some are also suitable for NUS data. Notable parametric methods include maximum likelihood (ML) [43–45], Bayesian [46], filter diagonalization method (FDM) [47], hybrid LP methods (LPSVD, HSVD, LPZ) [48–50], and generalized least-squares[51]. These methods all have in common a model consisting of a sum of exponentially decaying sinusoids. While they in principal are especially useful for characterizing spectra with minimal sampling requirements, they are less robust than regularization methods, prone to false positives and biased parameter values (frequencies, phases, amplitudes) when the number of sinusoids is underestimated, decay is not exponential, or the noise is not randomly distributed. Because the models only admit exponentially decaying sinusoids, spectra are frequently cosmetically reasonable even when biased.

The importance of computation and computer software for NMR spectroscopy can hardly be overstated. Richard Ernst put it rather succinctly in his contribution to a 1990 NATO workshop, "Without Computers – No Modern NMR" [52](Figure 3). Though the breadth of software applications today was largely envisioned at that workshop, breathtaking advances in computation – in speed, and lower cost – have enabled advances in the complexity of NMR spectra that can be routinely analyzed that exceed anything imagined in 1990. These improvements in computer hardware have fueled the creativity and audacity of software developers, to the point that hundreds of software packages are now routinely used in NMR labs and we are beginning to witness the emergence of an entirely new class of software:

meta-packages that link multiple different programs together to accomplish new and more complex tasks. Notable examples include ASDP [53], CS-Rosetta [54], COMPASS [55], and the NMRFAM Integrative NMR platform [56].

The Dark Side of Computing

There are downsides to the explosion in computer software for NMR. One is the complexity of the software environment for NMR, which makes it challenging to discover, install, and maintain a full range of applications. Comprehensive critical comparisons of software also become challenging, and with much of the software being developed by students in academic labs, lack of software persistence becomes a problem that in turn can diminish reproducibility. Another problem is somewhat counterintuitive: the staggering advances in computational speed can (seemingly) compensate for inefficient algorithm design. Not everything that can be computed today should be.

Relatively few software applications in NMR (or other scientific disciplines [57]) are developed by professional programmers or applied mathematicians. It is not surprising, then, that NMR software is dominated by heuristic algorithms, such as fixed-point methods (which iteratively apply the same operation until the resulting change is small) and greedy algorithms (which iteratively seek the local optimum). Examples of the former include iterative soft thresholding (IST) frequently used in CS, and of the latter include “best-first” approaches to automated assignment of nuclear Overhauser effects. The advantage of heuristic algorithms is they are usually easy to implement, but they typically lack formal underpinnings such as proof of the existence or uniqueness of solutions. Premature convergence is often a problem, because the step size can vanish before a fixed point or global optimum is reached [58].

Absent formal mathematical descriptions, peer review of software is difficult, and so it is up to the community to perform empirical critical comparisons. One promising approach is that of the “shared task”, employed by the Critical Assessment of Structure Prediction (CASP) [59] and related competitions, where a curated set of data is assembled and metrics for quantifying success are specified. This approach has been remarkably successful at not only identifying “best practices” but also in advancing the field of structure prediction. An alternative approach is to conduct a “matrix” comparison [60], in which methods are tested against all the data sets that were used as exemplars by the manuscripts describing the different methods. The rank order of the performance for different data sets can be revealing, and demonstrates that overall robustness is frequently neither correlated nor anti-correlated with optimum performance on a particular data set. While both these approaches to critical comparison are applicable in principle to the problem of spectrum analysis for NUS data, the lack of robust metrics (discussed below) for performance of nonlinear methods, as well as missing data and meta-data needed to reproduce results remain obstacles. Because so many modern methods lack formal mathematical descriptions, empirical critical comparisons are the only way to establish best practices, making it doubly important that software developers make available the complete software environment – including codes, data, and meta-data – that is required to reproduce their results. Resources such as the public data repository BMRB [61] and the NMR software platforms CCPN [62, 63], WeNMR [64], SBGrid [65],

and NMRbox [66] facilitate these comparisons, but it is incumbent upon journals, funding agencies, and the NMR community to ensure that published studies meet basic reporting and reproducibility standards.

Looking Forward

A cursory glance at the proceedings volume from the 1990 NATO workshop[67] reveals a pervasive sense of optimism that fully automated analysis of biomolecular NMR spectra was imminent. Two and a half decades later that optimism has yet to be fully rewarded, despite the fact that the speed of computation has increased 10^5 -fold, and the cost has decreased. The human analyst remains indispensable for biomolecular NMR studies.

In the narrower realm of spectrum analysis, however, the impact of improvements in computational speed have been profound. In the late 1980's, a 2D FFT required 12 hours on a VAX 11/780 "mini-supercomputer", mainly because of memory limitations. Although these 32-bit computers greatly facilitated programming by allowing direct addressing of large arrays, with memory prohibitively expensive few NMR labs could afford memory sufficient to hold even a modestly-size 2D data set in memory. The resulting "page faults" when transposing the data matrix meant that disk read/write speeds were limiting, not computational speed *per se*. With non-Fourier methods frequently requiring significantly more memory than the DFT (and 100-fold or more)computational effort), it's clear that non-Fourier methods would have found little application in NMR without advances in both computational speed and memory cost. Today a 2D DFT can be completed in milliseconds using an inexpensive laptop computer; a MaxEnt reconstruction can be completed in seconds.

The way forward to realizing higher sensitivity and resolution by leveraging the ability of non-Fourier methods to process NUS data is clear, as are the challenges. While the majority of NUS applications thus far have emphasized higher resolution at fixed measuring time or making 4D experiments feasible, the utility of NUS for enhancing sensitivity is receiving attention. Figure 4 gives an inkling of the sorts of gains that are already attainable. Further improvements in sampling schemes will enhance both sensitivity and resolution, but the field is in need of consensus quantitative metrics to drive advances. The DFT is linear, preserves norms, and treats signal and noise with equanimity. Not so with non-Fourier methods. Their nonlinearity renders SNR a poor proxy for sensitivity [68], but the challenge of finding robust quality metrics goes beyond nonlinearity, because of the different nature and magnitude of the nonlinearities of different methods. The challenge is to find a metric for sensitivity that is not only accurate despite nonlinearity, but is also *transferable*, making it possible to compare different methods. The magnitude of the challenge is clear by considering the "mock" data corresponding to the inverse DFT of the spectra produced by different non-Fourier methods: some methods stop when the χ^2 value for the match between the mock and the empirical data is comparable to the noise, but not smaller. Others match the mock data so that χ^2 is zero. For the former, peaks in the spectrum can be said to be statistically significant, but not for the latter.

Additional theoretical results from the CS and statistical communities can be expected to improve the concordance between theory and empirical results in NMR spectroscopy, and in turn lead to more robust and efficient spectral reconstruction algorithms. The concepts underlying NUS can be embraced more broadly, enabling NMR to stray from the constraints imposed by the Nyquist Theorem and the Jeener paradigm, to sparsely sample phase [69, 70] as well as time dimensions. One can imagine that sparse sampling approaches could productively engage with “single shot” methods [71] that sample space instead of time. Of course NUS concepts have long been used to advantage in imaging [72] for sampling spatial dimensions, and while there may be additional lessons from sparse sampling approaches in MRI that can be applied to spectroscopy, NMR spectra lack the strictly-positive nature of images and thus present unique challenges.

While the advent of powerful non-Fourier methods has unleashed this new wave of research on nonuniform sampling approaches, they have yet to reach their apotheosis. Beyond their use for processing NUS data, non-Fourier methods are also extremely useful for performing stable deconvolution [73, 74]. Although this capability has not been widely exploited, there is tantalizing anecdotal evidence that partial deconvolution of the natural lineshape [15, 75] can simultaneously enhance sensitivity and resolution. Additional resolution and sensitivity in multidimensional NMR spectroscopy afforded by non-Fourier methods not only promises to extend detection limits and the size and complexity of systems amenable to investigation using the most powerful ultra-high field instruments, but will, at virtually no cost and as a “retrofit”, also enhance older, less powerful spectrometers. There is indeed life Beyond Fourier, and it looks to be pretty exciting.

Acknowledgments

Work in my lab on signal processing in NMR has benefitted from the generous support of the US National Institutes of Health, via grants from the National Institute of General Medical Sciences (RR020125, GM102366, GM104517, GM111135). I am grateful for the advice and counsel of two brilliant mathematicians, Alan Stern and David Donoho, who have always and generously explained the hard bits to me. I am indebted to my colleagues Hari Arthanari, Jonathan Helmus, Sven Hyberts, Mark Maciejewski, Mehdi Mobli, Hafez Monajemi, Tatyana Polenova, David Rovnyak, Julian Rujan, Peter Schmieder, and Adam Schuyler for fruitful collaborations. I thank Clemens Anklin, Andy Byrd, George Levy, and Marc-André Delsuc for discussions on the provenance of LP extrapolation. Soham Kale and Irina Bezsonova graciously supplied the data for Figure 4. Finally, I am especially grateful for decades of advice and encouragement from Gerhard Wagner, whose vision and courage continue to inspire.

References

1. The Human Dark Proteome Initiative. 2015. <http://darkproteome.org/>
2. Wright PE, Dyson HJ. Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol.* 2015; 16:18–29. [PubMed: 25531225]
3. Kaptein R, Dijkstra K, Nicolay K. Laser photo-CIDNP as a surface probe for proteins in solution. *Nature.* 1978; 274:293–294. [PubMed: 683312]
4. Mueller RM, Buchal C, Folle HR, Kubota M, Powell F. A Double-stage Nuclear Demagnetization Refrigerator. *Cryogenics.* 1980; 20:395–407.
5. Kenyon B, Kleinberg R, Straley C, Gubelin G, Morriss C. Nuclear Magnetic Resonance Imaging — Technology for the 21st Century. *Oilfield Review.* 1995; 7:19–33.
6. Stark RE, Haberkorn RA, Griffin RE. ¹⁴N NMR Determination of NH Bond Lengths in Solids. *J Chem Phys.* 1978; 68:1996–1997.

7. Roberts JE, Harbison GS, Munowitz MG, Herzfeld J, Griffin RG. Measurement of Heteronuclear Bond Distances in Polycrystalline Solids by Solid-State NMR Techniques. *J Am Chem Soc.* 1987; 109:4163–4169.
8. Lovchinsky I, Sushkov AO, Urbach E, de Leon NP, Choi S, De Greve K, Evans R, Gertner R, Bersin E, Müller C, McGuinness L, Jelezko F, Walsworth RL, Park H, Lukin MD. Nuclear magnetic resonance detection and spectroscopy of single proteins using quantum logic. *Science.* 2016; 351:836–841. [PubMed: 26847544]
9. Ernst RR, Anderson WA. Application of Fourier Transform Spectroscopy to Magnetic Resonance. *Rev Sci Instrum.* 1966; 37:93–102.
10. Cooley JW, Tukey JW. An algorithm for the machine calculation of complex Fourier series. *Math Comput.* 1965; 19:297–301.
11. Middleton D, Van Vleck JH. Theoretical Comparison of the Visual A Aural and Meter Reception of Pulsed Signals in the Presence of Noise. *J Appl Phys.* 1946; 17:940–971.
12. Wiener N. Generalized harmonic analysis. *Acta Math.* 1930:117–258.
13. Jerison, D., Singer, IM., Stroock, DW. The Legacy of Norbert Wiener: A Centennial Symposium; American Mathematical Society; 1997. p. 95
14. Stern AS, Li KB, Hoch JC. Modern spectrum analysis in multidimensional NMR spectroscopy: comparison of linear-prediction extrapolation and maximum-entropy reconstruction. *J Am Chem Soc.* 2002; 124:1982–1993. [PubMed: 11866612]
15. Sibisi S, Skilling J, Brereton RG, Laue ED, Staunton J. Maximum entropy signal processing in practical NMR spectroscopy. *Nature.* 1984; 311:446–447.
16. Shannon CE. A Mathematical Theory of Communication. 1948:379–423.
17. Burg, JP. Maximum entropy spectrum analysis. Stanford University; 1975.
18. Jaynes, ET. Where do we Stand on Maximum Entropy?. In: Levine, RD., Tribus, M., editors. The Maximum Entropy Formalism. MIT Press; Cambridge: 1979. p. 15
19. Skilling J, Bryan R. Maximum entropy image reconstruction: general algorithm. *Mon Not R Astron Soc.* 1984; 211:111–124.
20. Hoch JC, Stern AS, Donoho DL, Johnstone IM. Maximum Entropy Reconstruction of Complex (Phase-Sensitive) Spectra. *J Magn Reson.* 1990; 86:236–246.
21. Daniell GJ, Hore PJ. Maximum Entropy and NMR - A new Approach. *J Magn Reson.* 1989; 84:515–536.
22. Barna JCJ, Laue ED. Conventional and exponential sampling for 2D NMR experiments with application to a 2D NMR spectrum of a protein. *J Magn Reson.* 1987; 75:384–389.
23. Barna JCJ, Laue ED, Mayger MR, Skilling J, Worrall SJP. Exponential sampling: an alternative method for sampling in two dimensional NMR experiments. *J Magn Reson.* 1987; 73:69–77.
24. Kup e , Freeman R. Projection-reconstruction technique for speeding up multidimensional NMR spectroscopy. *J Am Chem Soc.* 2004; 126:6429–6440. [PubMed: 15149240]
25. Kup e E, Freeman R. The Radon Transform: A new scheme for fast multidimensional NMR. *Concepts Magn Reson.* 2004; 22A:4–11.
26. Hiller S, Fiorito F, Wüthrich K. Automated projection spectroscopy (APSY). *Proc Natl Acad Sci USA.* 2005; 102:10876–10888. [PubMed: 16043707]
27. Eghbalnia HR, Bahrami A, Tonelli M, Hallenga K, Markley JL. High-resolution iterative frequency identification for NMR as a general strategy for multidimensional data collection. *J Am Chem Soc.* 2005; 127:12528–12536. [PubMed: 16144400]
28. Mobli M, Stern AS, Hoch JC. Spectral reconstruction methods in fast NMR: Reduced dimensionality, random sampling and maximum entropy. *J Magn Reson.* 2006
29. Coggins BE, Zhou P. Polar Fourier transforms of radially sampled NMR data. *J Magn Reson.* 2006; 182:84–95. [PubMed: 16820311]
30. Kazimierczuk K, Orekhov V. Non-uniform sampling: post-Fourier era of NMR data collection and processing. *Magn Reson Chem.* 2015; 53:921–926. [PubMed: 26290057]
31. Mobli M, Maciejewski MW, Schuyler AD, Stern AS, Hoch JC. Sparse sampling methods in multidimensional NMR. *Phys Chem Chem Phys.* 2012; 14:10835–10843. [PubMed: 22481242]

32. Hyberts SG, Arthanari H, Robson SA, Wagner G. Perspectives in magnetic resonance: NMR in the post-FFT era. *Journal of Magnetic Resonance*. 2014; 241:60–73. [PubMed: 24656081]
33. Logan, B. Properties of high-pass signals. Department of Electrical Engineering, Columbia University; 1965.
34. Donoho DL, Logan BF. Signal Recovery and the Large Sieve. *SIAM Journal on Applied Mathematics*. 1992; 52:577–591.
35. Donoho DL. Compressed Sensing. *IEEE Transactions on Information Theory*. 2006; 52:1289–1306.
36. Candès EJ, Romberg JK, Tao T. Stable signal recovery from incomplete and inaccurate measurements. *Comm Pure Appl Math*. 2006; 59:1207–1223.
37. Donoho DL, Johnstone IM, Stern AS, Hoch JC. Maximum Entropy and the Nearly Black Object (with discussion). *J Roy Stat Soc B*. 1992; 54:41–81.
38. Monajemi H, Jafarpour S, Gavish M, Collaboration SC, Donoho DL. Deterministic matrices matching the compressed sensing phase transitions of Gaussian random matrices. *Proc Nat Acad Sci USA*. 2013; 110:1181–1186. [PubMed: 23277588]
39. Donoho DL, Tanner J. Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing. *Phil Trans R Soc*. 2009; 367:4273–4293.
40. Stern AS, Donoho DL, Hoch JC. NMR data processing using iterative thresholding and minimum $l(1)$ -norm reconstruction. *J Magn Reson*. 2007; 188:295–300. [PubMed: 17723313]
41. van Cittert PH. Zum Einfluss der Spaltbreite auf die Intensitätsverteilung in Spektrallinien. II. *Zeitschrift für Physik*. 1931; 69:298–308.
42. Monajemi, H. Statistics. Stanford University; 2016. Phase Transitions in Deterministic Compressed Sensing, with Application to Magnetic Resonance Spectroscopy.
43. Chylla RA, Volkman BF, Markley JL. Practical model fitting approaches to the direct extraction of NMR parameters simultaneously from all dimensions of multidimensional NMR spectra. *J Biomol NMR*. 1998; 12:277–297. [PubMed: 9751999]
44. Chylla RA, Markley JL. Theory and application of the maximum likelihood principle to NMR parameter estimation of multidimensional NMR data. *J Biomol NMR*. 1995; 5:245–258. [PubMed: 7787422]
45. Miller MI, Greene AS. Maximum-likelihood estimation for nuclear magnetic resonance spectroscopy. *Journal of Magnetic Resonance* (1969). 1989; 83:525–548.
46. Bretthorst GL. Bayesian Analysis I. Parameter estimation using quadrature NMR models. *J Magn Reson*. 1990; 88:533–551.
47. Mandelshtam VA, Taylor HS, Shaka AJ. Application of the filter diagonalization method to one- and two-dimensional NMR spectra. *J Magn Reson*. 1998; 133:304–312. [PubMed: 9716473]
48. Barkhuijsen H, Beer Rd, Bovee WM, Creyghton JH, Ormondt Dv. Application of linear prediction and singular value decomposition (LPSVD) to determine NMR frequencies and intensities from the FID. *Magn Reson Med*. 1985; 2:86–89. [PubMed: 3831681]
49. de Beer R, van den Boogaart A, van Ormondt D, Pijnappel WW, den Hollander JA, Marien AJ, Luyten PR. Application of time-domain fitting in the quantification of in vivo ^1H spectroscopic imaging data sets. *NMR Biomed*. 1992; 5:171–178. [PubMed: 1449952]
50. Tang J, Lin CP, Bowman MK, Norris JR. An alternative to fourier transform spectral analysis with improved resolution. *Journal of Magnetic Resonance* (1969). 1985; 62:167–171.
51. Hua Y. The most efficient implementation of the IQML algorithm. *IEEE Trans Sig Proc*. 1994; 42:2203–2204.
52. Ernst, RR. Without Computers — No Modern NMR. In: Hoch, JC, Poulsen, FM., Redfield, C., editors. *Computational Aspects of the Study of Biological Macromolecules by Nuclear Magnetic Resonance Spectroscopy*. Plenum Press; New York: 1991. p. 1-25.
53. Huang YJ, Tejero R, Powers R, Montelione GT. A topology-constrained distance network algorithm for protein structure determination from NOESY data. *Proteins*. 2006; 62:587–603. [PubMed: 16374783]

54. Shen Y, Vernon R, Baker D, Bax A. De novo protein structure generation from incomplete chemical shift assignments. *J Biomol NMR*. 2009; 43:63–78. [PubMed: 19034676]
55. Courtney JM, Ye Q, Nesbitt AE, Tang M, Tuttle MD, Watt ED, Nuzzio KM, Sperling LJ, Comellas G, Peterson JR, Morrissey JH, Rienstra CM. Experimental Protein Structure Verification by Scoring with a Single, Unassigned NMR Spectrum. *Structure*. 2015; 23:1958–1966. [PubMed: 26365800]
56. Lee W, Cornilescu G, Dashti H, Eghbalnia HR, Tonelli M, Westler WM, Butcher SE, Henzler-Wildman KA, Markley JL. Integrative NMR for biomolecular research. *J Biomol NMR*. 2016; 64:307–332. [PubMed: 27023095]
57. Chang, K. Simons Foundation Looks to Enhance Computation Tools for Science. *The New York Times*; New York: Nov 23. 2016
58. Stern AS, Donoho DL, Hoch JC. NMR data processing using iterative thresholding and minimum l1-norm reconstruction. *Journal of Magnetic Resonance*. 2007; 188:295–300. [PubMed: 17723313]
59. Moulton J, Fidelis K, Kryshchuk A, Schwede T, Tramontano A. Critical assessment of methods of protein structure prediction (CASP)--round x. *Proteins*. 2014; 82(Suppl 2):1–6.
60. Bernau C, Riestler M, Boulesteix AL, Parmigiani G, Huttenhower C, Waldron L, Trippa L. Cross-study validation for the assessment of prediction algorithms. *Bioinformatics*. 2014; 30:i105–112. [PubMed: 24931973]
61. Ulrich EL, Akutsu H, Doreleijers JF, Harano Y, Ioannidis YE, Lin J, Livny M, Mading S, Maziuk D, Miller Z, Nakatani E, Schulte CF, Tolmie DE, Kent Wenger R, Yao H, Markley JL. BioMagResBank. *Nucleic Acids Res*. 2008; 36:D402–408. [PubMed: 17984079]
62. Skinner SP, Fogh RH, Boucher W, Ragan TJ, Mureddu LG, Vuister GW. CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *J Biomol NMR*. 2016; 66:111–124. [PubMed: 27663422]
63. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*. 2005; 59:687–696. [PubMed: 15815974]
64. van der Schot G, Bonvin AM. Performance of the WeNMR CS-Rosetta3 web server in CASD-NMR. *J Biomol NMR*. 2015; 62:497–502. [PubMed: 25982706]
65. Morin A, Eisenbraun B, Key J, Sanschagrin PC, Timony MA, Ottaviano M, Sliz P. Collaboration gets the most out of software. *Elife*. 2013; 2:e01456. [PubMed: 24040512]
66. Maciejewski MW, Schuyler AD, Gryk MR, Moraru II, Romero PR, Ulrich EL, Eghbalnia HR, Livny M, Delaglio F, Hoch JC. NMRbox: A Resource for Biomolecular NMR Computation. *Biophys J*. 2017 in press.
67. Computational Aspects of the Study of Biological Macromolecules by Nuclear Magnetic Resonance Spectroscopy. Plenum Press; New York: 1991.
68. Donoho DL, Johnstone IM, Stern AS, Hoch JC. Does the maximum entropy method improve sensitivity? *Proc Natl Acad Sci U S A*. 1990; 87:5066–5068. [PubMed: 11607089]
69. Maciejewski MW, Fenwick M, Schuyler AD, Stern AS, Gorbatyuk V, Hoch JC. Random phase detection in multidimensional NMR. *Proc Natl Acad Sci U S A*. 2011; 108:16640–16644. [PubMed: 21949370]
70. Bostock MJ, Holland DJ, Nietlispach D. Improving resolution in multidimensional NMR using random quadrature detection with compressed sensing reconstruction. *J Biomol NMR*. 2016
71. Frydman L, Scherf T, Lupulescu A. The acquisition of multidimensional NMR spectra within a single scan. *Proc Natl Acad Sci USA*. 2002; 99:15858–15862. [PubMed: 12461169]
72. Lustig M, Donoho D, Pauly JM. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn Reson Med*. 2007; 58:1182–1195. [PubMed: 17969013]
73. Shimba N, Stern AS, Craik CS, Hoch JC, Dötsch V. Elimination of ¹³C α splitting in protein NMR spectra by deconvolution with maximum entropy reconstruction. *J Am Chem Soc*. 2003; 125:2382–2383. [PubMed: 12603112]
74. Jordan JB, Kovacs H, Wang Y, Mobli M, Luo R, Anklin C, Hoch JC, Kriwacki RW. Three-dimensional ¹³C-detected CH₃-TOCSY using selectively protonated proteins: facile methyl

resonance assignment and protein structure determination. *J Am Chem Soc.* 2006; 128:9119–9128. [PubMed: 16834385]

75. Ni F, Levy GC, Scheraga HA. Simultaneous resolution enhancement and noise suppression in NMR signal processing by combined use of maximum entropy and fourier self-deconvolution methods. *J Magn Reson.* 1986; 66:385–390.

Highlights

- Non-Fourier methods have a rich history in science and engineering, but were slow to gain traction in NMR spectroscopy.
- Non-Fourier and nonuniform sampling (NUS) approaches are practical means of achieving very high resolution in the indirect dimensions of 3+ dimensional experiments.
- Reductions of measuring time of an order of magnitude are easily attained for 3D experiments.
- Gain in sensitivity per unit measuring time comparable to or exceeding that provided by cryogenic probes is attainable.
- Recent theoretical and algorithmic advances pave the way for additional improvements in spectral quality.

Figure 1. Statisticians having unusually weighty impact on NMR signal processing. Left: Benjamin F. Logan (photo by John Byrne Cooke, with permission) Right: David Donoho, in front of the Russell H. Varian Laboratory of Physics on the Stanford University Campus.

Figure 2.

Donoho-Tanner “phase transitions” in compressed sensing, adapted from Monajemi [42]. The color corresponds to the probability of accurate recovery of the spectrum from sparse (nonuniformly sampled) data. Blue is low probability (failure), red is high probability (success). The horizontal axis reports sampling coverage, which is the number of sampled data values divided by the number of spectrum values to be recovered. The vertical axis reports the signal sparsity, the number of non-zero values in the spectrum divided by the total number elements in the spectrum. The vertical columns denoted by rectangular boxes depict a fixed number of samples at low (magenta) and high magnetic field, and fixed digital resolution. At higher field, the spectrum becomes more sparse as the “space” between signals grows, however additional samples are required due to the increased sampling rate necessitated by greater spectral dispersion. The effects are nearly compensatory if the linewidths of the signals are unchanged, so transitioning from an experiment that resides in the failure region at low field to higher field traverses a path (yellow arrow) that is nearly parallel to the Donoho-Tanner phase transition separating failure from success. If the linewidths become narrower at higher field, for example due to TROSY effects, the path becomes steeper (black arrow), and the potential for crossing the phase boundary into the success region becomes greater.

Figure 3.

Polaroid (SX70) photographs from the 1990 NATO Advanced Research Workshop “Computational Aspects of the Study of Biological Macromolecules by NMR”, Barga, Italy. Top row, left: Tom James, Frank Delaglio and Ed Olejniczak are visible in the foreground. Center: Christina Redfield (foreground). Ruud Scheek and Christian Griesinger are visible in the background, facing the camera. Right: Werner Braun, Martin Billeter, (both bearded) Tad Holak, Ruud Scheek. Bottom row, left: Angela Gronenborn. Center: Martin Billeter, Dennis Hare*, Mogens Kjaer, Flemming Poulsen* standing at left. Frank Delaglio and Jim Prestegard sitting at right. Right: Frank Delaglio, Ed Olejniczak (facing the camera in the background), John Markley and Philip Bolton, foreground. [* Dennis Hare passed away in 2007. He was noted in his island community for ferrying cancer patients in his Cessna 182 for treatment. Flemming Poulsen died in 2011. He was the patriarch of bio-NMR in Denmark, having established two high-field facilities: first at the Carlsberg Lab, later at the University of Copenhagen. Both Dennis and Flemming made important contributions to NMR data processing and analysis.]

Figure 4.

2D and 1D cross-sections through a plane of the 3D HNCACB spectra for the UBL3 domain of USP7. Top: Using data collected using conventional uniform sampling, DFT processing with linear prediction extrapolation and shifted sine-squared apodization in both indirect dimensions. The 1D cross-section is taken at the ^{15}N frequency indicated by the dashed line. Bottom: Using data collected using nonuniform sampling exponentially biased to shorter evolution times and processed using MaxEnt reconstruction in both indirect dimensions. Each experiment required 48 hours of measurement time. The NUS coverage was 25% of the evolution times sampled in the US experiment; the factor of four was used to increase the number of transients acquired at each evolution time by a factor of four.