



Published in final edited form as:

*Magn Reson Chem.* 2015 November ; 53(11): 908–912. doi:10.1002/mrc.4287.

## A New Approach to Compressed Sensing for NMR

Alan S. Stern<sup>1</sup> and Jeffrey C. Hoch<sup>2,\*</sup>

<sup>1</sup>Rowland Institute at Harvard, 100 Edwin H. Land Blvd. Cambridge, MA 02142

<sup>2</sup>UConn Health, Department of Molecular Biology and Biophysics, 263 Farmington Ave. Farmington, CT 06030-3305

### Abstract

Compressed sensing (CS) has attracted a great deal of recent interest as an approach for spectrum analysis of nonuniformly sampled NMR data. Although theoretical justification for the method is abundant, it suffers from several weaknesses, among them poor convergence of some algorithms, and it remains an open question whether NMR spectra satisfy the sparsity requirements of CS theorems. The versions of CS used in NMR involve minimizing the  $l_1$  norm of the spectrum. They bear similarity to maximum entropy (MaxEnt) reconstruction, but critical comparison of the methods can be difficult. Here we describe a formalism that places CS and MaxEnt reconstruction on equal footing, enabling critical comparison of the two methods. We also describe a new algorithm for CS that restricts the computation of the  $l_1$  norm to the real channel for complex spectra and ensures causality. Preliminary 1D results demonstrate that this approach ameliorates some artifacts that can occur when using the  $l_1$  norm of the complex spectrum.

---

Spectrum analysis of time series data has captured the attention of mathematicians, scientists, and engineers for centuries. NMR spectroscopists are relative newcomers, compared with illustrious figures such as Gauss(1), Prony(2), Fourier(3), and Tukey(4). NMR spectroscopists continue to mine the rich literature on spectrum analysis for "novel" methods, a task sometimes made difficult because of terminology employed by different fields: identical or closely related methods are often described in very different terms that obscure relationships among the methods.

The approach known as compressed sensing (CS), frequently attributed to Candes(5), is finding widespread application for spectrum analysis of NMR data collected using nonuniform sampling (NUS)(6–11). CS methods typically work by minimizing the  $l_1$  norm of the spectrum (i.e. the sum of its absolute values). Logan's Theorem(12) states that perfect recovery of the spectrum is possible even from incomplete and noisy measurements by minimizing the  $l_1$  norm of the spectrum while ensuring consistency with the measured data, provided that certain conditions on sparsity and noise level are met, and the spectrum is band-limited.

---

\*Correspondence to hoch@uchc.edu.

#### Reproducible research

Software employed in this study is available from <http://rnmrtk.uchc.edu>. Scripts and sampling schedules used to produce the figures are available as supplementary data.

Common CS algorithms involve iteratively thresholding the spectrum(10, 13– 15). This approach has a long and rich prior history, including iterative thresholding image processing algorithms due to van Cittert(16), Jansson(17), and Papoulis(18). Iterative thresholding algorithms for NMR spectra involve two steps. In the thresholding step, values in the frequency spectrum are set to zero if they are below some threshold in absolute value. Then in the replacement step, the inverse Fourier transform of the thresholded spectrum is computed. Values corresponding to measured data are then replaced by the measured values, and the result is Fourier transformed to the frequency domain. These two steps are repeated alternately until convergence. The connection between the iterative thresholding algorithms and the theoretical work of Logan was made explicit by Stern et al.(15), who proved that the fixed-point of iterative *soft* thresholding (IST) has minimum  $l_1$  norm. Iterative soft thresholding differs from iterative thresholding only in the thresholding step: values below the threshold are set to zero, and values above the threshold are reduced in absolute value by the threshold amount (keeping the same phase).

While iterative thresholding algorithms hold strong heuristic appeal and are trivial to code, as fixed-point methods they are sometimes subject to premature convergence(15), in which the step size becomes vanishingly small before the optimum (minimum  $l_1$  norm) is reached. A potentially more fundamental limitation of CS is that NMR spectra may not meet the sparsity conditions of Logan's Theorem or other CS theorems(19). The Lorentzian lineshapes that occur in NMR spectroscopy are not band-limited, and furthermore are not localized to single points in the frequency domain. A difference between NMR spectra and imaging applications of CS is that images are non-negative and real, whereas NMR spectra are complex (and even their absorptive components need not be non-negative). Finally, CS methods are not guaranteed to produce a unique spectrum.

CS has been proposed as an alternative to maximum entropy (MaxEnt) reconstruction for spectrum analysis of NUS NMR data(6–11). Critical comparison of the two methods is not straightforward, however, because most of the proposed algorithms constrain the inverse Fourier transform of the spectrum to exactly match the measured data(9, 10) (for example by stopping the iterative thresholding procedure following a replacement step). It is important to note that the proof of the equivalence between IST and  $l_1$ -norm minimization(15) holds only when the final step is thresholding, not replacement, and the results of these algorithms cannot accurately be said to minimize the  $l_1$  norm. MaxEnt reconstruction as originally formulated by Wernecke and D'Addario(20), in contrast, merely ensures that the inverse transform of the spectrum is statistically consistent with the measured data to within a prior estimate of the noise. While other MaxEnt methods such as Forward Maximum Entropy (FM)(21) and Maximum Entropy Interpolation (MINT)(22) that exactly or closely match the measured data have been demonstrated and offer an advantage of being nearly linear, they "over-fit" the measured data.

Here we describe an approach to CS that constrains the spectrum to agree with the empirical data in a manner analogous to MaxEnt reconstruction. By ensuring a known level of agreement between the reconstructed spectra and the known (empirical) data, this enables fair comparisons of MaxEnt and minimum  $l_1$  norm spectra that have the same level of agreement. The approach employs robust gradient search and a metric for measuring the

degree of convergence that is more strict than relying on the step size. We restrict the computation of the  $l_1$  norm to the real component of complex spectra. Provided that the phase of the spectrum can be determined and adjusted so that the real component is absorptive, this helps limit the support of Lorentzian peaks to relatively few frequency values, thus improving the overall sparsity compared to a computation that includes dispersive imaginary components, an idea explored by Mayzel et al.(23). In addition, we incorporate support for including a convolution kernel in the Fourier inversion between the spectral estimate and the time domain, enabling deconvolution to help narrow spectral peaks as an alternative method for improving the sparsity.

## Theory

Our approach to  $l_1$  norm regularization is based on a formalization similar to that of MaxEnt regularization(24), although there are a few differences in the details below. The desired spectrum  $\mathbf{f}$  is one which minimizes a regularization functional  $S(\mathbf{f})$  subject to two constraints:  $\mathbf{f}$  must be causal and  $\mathbf{f}$  must be consistent with the measured data.  $S(\mathbf{f})$  is defined to be the  $l_1$  norm computed for just the real component of  $\mathbf{f}$  (or real-real quadrant for a two-dimensional spectrum, and so on):

$$S(\mathbf{f}) = \sum_{n=0}^{N-1} |\text{real}(f_n)|, \quad (1)$$

where  $N$  is the number of points in the spectrum. The data constraint is defined in terms of the root-mean-square (RMS) difference between the inverse discrete Fourier transform (IDFT) of  $\mathbf{f}$  and the measured data  $\mathbf{d}$ :

$$C(\mathbf{f}, \mathbf{d}) = \sqrt{\sum_m |\text{IDFT}(\mathbf{f})_m - d_m|^2 / nc} \quad (2)$$

Here the sum runs over only the measured data points rather than all  $N$  points in  $\text{IDFT}(\mathbf{f})$ , and  $nc$  is the total number of components included in the sum (real and imaginary parts of the data are counted as separate components). Eq. (2) can be generalized to support deconvolution of a known line shape from the measured data by factoring in a convolution kernel  $\mathbf{k}$ :

$$C(\mathbf{f}, \mathbf{d}) = \sqrt{\sum_m |k_m \text{IDFT}(\mathbf{f})_m - d_m|^2 / nc} \quad (3)$$

Thus, we seek to find  $\mathbf{f}$  such that  $S(\mathbf{f})$  is minimal subject to the constraints that  $\mathbf{f}$  is causal and

$$C(\mathbf{f}, \mathbf{d}) \leq \text{aim}, \quad (4)$$

where  $aim$  is an estimate of the RMS noise level in  $\mathbf{d}$ . Because the  $S$  and  $C$  functionals are both convex, the existence of a minimum is guaranteed, and the minimum will satisfy either  $C(\mathbf{f}, \mathbf{d}) = aim$  or else  $\mathbf{f} = 0$ . However, because  $S$  is not *strictly* convex, the minimum need not be unique.

It is well known that for causal spectra, the imaginary part is determined by the real part via the Hilbert transform(25). Therefore there is no need to store the imaginary part of  $\mathbf{f}$  during the computation, and indeed our algorithm stores only the real part, which we will call  $\mathbf{h}$ :

$$h_n = \text{real}(f_n). \quad (5)$$

The problem of computing  $\text{IDFT}(\mathbf{f})$  from  $\mathbf{h}$  is solved as follows. The real part of  $f_n$  is one-half the sum of  $f_n$  and its complex conjugate, and complex conjugation in the frequency domain corresponds to complex conjugation plus reversal in the time domain(24). Therefore

$$\text{IDFT}(\mathbf{h})_m = (1/2)[\text{IDFT}(\mathbf{f})_m + \text{conj}(\text{IDFT}(\mathbf{f})_{N-m})] \quad (6)$$

(where we take  $\text{IDFT}(\mathbf{f})_N$  to be the same as  $\text{IDFT}(\mathbf{f})_0$ ). Causality says that in the time domain, the data points at negative times are all equal to 0; by the cyclic nature of the discrete Fourier transform this means that  $\text{IDFT}(\mathbf{f})_m$  is equal to 0 for all  $m > N/2$ . Hence from Eq. (6) it follows that:

$$\text{IDFT}(\mathbf{h})_m = \begin{cases} \text{real}(\text{IDFT}(\mathbf{f})_m) & \text{if } m=0, \\ \text{IDFT}(\mathbf{f})_{m/2} & \text{if } 1 \leq m < N/2. \end{cases} \quad (7)$$

Although this leaves the imaginary part of  $\text{IDFT}(\mathbf{f})_0$  undetermined, the rest of the terms appearing in Eqs. (2) or (3) can be gotten from  $\text{IDFT}(\mathbf{h})$ :

$$\text{real}(\text{IDFT}(\mathbf{f})_0) = \text{IDFT}(\mathbf{h})_0, \quad (8)$$

$$\text{IDFT}(\mathbf{f})_m = 2 \text{IDFT}(\mathbf{h})_m \text{ for } 1 \leq m < N/2. \quad (9)$$

(We require that  $N$  be more than twice as large as the largest index  $m$  appearing in the measured data, so that the sums in Eqs. (2) and (3) run over values of  $m$  that are all less than  $N/2$ . This is tantamount to zero-filling the measured data by at least a factor of two.)

In addition, it is generally recognized that owing to the difference between the continuous and discrete Fourier transforms, in order to obtain an accurate base line it is necessary to divide the first point of the time-domain data by two prior to transformation(26). To put it another way, the  $m=0$  term in the definition of  $C$  should be computed from the difference between  $2 \times \text{IDFT}(\mathbf{f})_0$  and  $d_0$  rather than the difference between  $\text{IDFT}(\mathbf{f})_0$  and  $d_0$ . Note: Since

the imaginary part of  $\text{IDFT}(\mathbf{f})_0$  is not determined by  $\mathbf{h}$ , its component must be omitted from the sums in Eqs. (2) or (3).

Putting these ideas together, we reformulate the problem: We seek a real spectrum  $\mathbf{h}$  such that  $S(\mathbf{h})$  is minimal subject to the constraint that  $C(\mathbf{h}, \mathbf{d}) = aim$ , where  $C(\mathbf{h}, \mathbf{d})$  is now given by:

$$C(\mathbf{h}, \mathbf{d}) = \sqrt{\sum_m |2k_m \text{IDFT}(\mathbf{h})_m - d_m|^2 / nc} \quad (10)$$

(and the imaginary component is not included in the  $m=0$  term of the sum).

This constrained minimization problem can be converted to an unconstrained minimization by introducing a Lagrange multiplier  $\lambda$ : Find  $\mathbf{h}$  minimizing  $S(\mathbf{h}) + \lambda C(\mathbf{h}, \mathbf{d})$ , where  $\lambda$  is selected so that the minimum satisfies  $C(\mathbf{h}, \mathbf{d}) = aim$ . However, we choose instead to set  $\tau$  equal to  $1/\lambda$  and find  $\mathbf{h}$  minimizing the objective functional

$$Q(\mathbf{h}, \mathbf{d}, \tau) = \tau S(\mathbf{h}) + C(\mathbf{h}, \mathbf{d}). \quad (11)$$

Using  $\tau$  rather than  $\lambda$  makes no difference to the final result, but it does permit a simple comparison with the IST algorithm because  $\tau$  turns out to be equal to the threshold value(15).

(Our algorithm does not constrain the values of  $\mathbf{h}$  to be non-negative, but such a constraint would be very easy to add. Note that a major difference between NMR and most imaging application of CS is the absence of a non-negativity constraint.)

The algorithm employs a slightly modified Conjugate Gradient (CG) minimization(27) to find  $\mathbf{h}$ . The algorithm can be used in two different modes. In the first mode, the user specifies  $\tau$  beforehand and the algorithm simply computes  $\mathbf{h}$  so as to minimize  $Q(\mathbf{h}, \mathbf{d}, \tau)$ . This mode is analogous to the “constant- $\lambda$ ” algorithm we previously described for MaxEnt reconstruction (28) and thus can be called the “constant- $\tau$ ” mode. This is useful when computing 2D spectra for successive planes of a 3D data set (or 1D spectra from a 2D data set), as it ensures that nonlinearities are similar for different planes and thus simplifies calibration of the nonlinearities for quantitative applications. Applied to 3D data, a typical plane or set of planes would be selected and iteratively reconstructed using different values of  $\tau$ , and the entire set of planes is then reconstructed using the value of  $\tau$  that yields  $C(\mathbf{h}, \mathbf{d})$  with the desired value of  $aim$ . In the second mode, the user specifies  $aim$  and the algorithm finds both  $\tau$  and  $\mathbf{h}$  so that  $\mathbf{h}$  minimizes  $Q(\mathbf{h}, \mathbf{d}, \tau)$  and  $C(\mathbf{h}, \mathbf{d}) = aim$ . This is the “constant- $aim$ ” mode. Applied to planes of a 3D data set, this mode typically yields different values for  $\tau$ , which is an obstacle to quantitation.

CG minimization is quite efficient when the function to be minimized is quadratic. In our case  $C(\mathbf{h}, \mathbf{d})$  is quadratic, and  $S(\mathbf{h})$  is linear *provided* the computation is restricted to a domain in which none of the  $h_n$  values change sign. Indeed, the absolute value function is

not differentiable at the origin, which causes difficulties for gradient-based techniques. Our algorithm handles this situation by adopting the following conventions at indices  $n$  for which  $h_n = 0$ :

If  $|\nabla C(\mathbf{h}, \mathbf{d})_n| \leq \tau$  then  $\nabla S(\mathbf{h})_n$  and  $\nabla Q(\mathbf{h}, \mathbf{d}, \tau)_n$  are both set to 0. Under these circumstances, any small change to  $h_n$  will result in an increase in  $Q$ ; setting  $\nabla Q(\mathbf{h}, \mathbf{d}, \tau)_n$  to 0 prevents the algorithm from changing  $h_n$ .

Otherwise  $\nabla S(\mathbf{h})_n$  is set to either 1 or  $-1$  according to whether  $\nabla C(\mathbf{h}, \mathbf{d})_n$  is negative or positive. Under these circumstances,  $Q$  can be decreased by changing  $h_n$  in the direction opposite  $\nabla C(\mathbf{h}, \mathbf{d})_n$  and after such a change,  $\nabla S(\mathbf{h})_n$  will be equal to the value adopted here (1 or  $-1$ ).

During each step of the CG minimization, if the step would cause any element  $h_n$  to change sign,  $h_n$  is set to 0 instead. This prevents the quadratic extrapolation used by CG from being applied outside its domain of validity. In addition, the CG computation is reset (by forgetting the previous steps) whenever the support set of  $\nabla S(\mathbf{h})$  (i.e., the set of indices  $n$  at which  $\nabla S(\mathbf{h})_n$  is nonzero) changes; the next step will follow the gradient of  $Q$  rather than the conjugate gradient. It is expected that as the computation converges toward the final answer, these exceptional cases will arise relatively rarely.

The algorithm computes a test statistic before each step:

$$\text{test} = \tau^{-1} \sqrt{\sum_{n=0}^{N-1} (\nabla Q(\mathbf{h}, \mathbf{d}, \tau))^2 / \text{nsupp}} \quad (12)$$

where  $\text{nsupp}$  is the size of the support set of  $\nabla S(\mathbf{h})$  ( $\nabla Q(\mathbf{h})_n$  is necessarily equal to 0 at each  $n$  outside the support set). This statistic measures the degree of convergence. In the first mode of operation, the minimization ends when  $\text{test} < 0.0001$ .

The second mode is more complicated. The algorithm makes an initial overestimate of  $\tau$ , which it decreases over time. As long as  $C(\mathbf{h}, \mathbf{d})$  remains above  $\text{aim}$ , the algorithm decreases  $\tau$  by 0.8 and restarts the CG minimization whenever  $\text{test}$  falls below 0.01. But once  $C(\mathbf{h}, \mathbf{d})$  has fallen below  $\text{aim}$ , the algorithm changes strategy. Now whenever  $\text{test}$  is below 0.01 and is also below  $|C(\mathbf{h}, \mathbf{d}) - \text{aim}| / \text{aim}$ , the algorithm takes a single step along  $\nabla C(\mathbf{h}, \mathbf{d})$ . The length of the step is chosen to make the new value of  $C(\mathbf{h}, \mathbf{d})$  equal to  $\text{aim}$ , and following the step,  $\tau$  is recalculated as:

$$\tau = - \sum_{n=0}^{N-1} \nabla C(\mathbf{h}, \mathbf{d})_n \nabla S(\mathbf{h})_n / \text{nsupp} \quad (13)$$

At all other iterations the algorithm uses CG minimization to decrease  $Q(\mathbf{h}, \mathbf{d}, \tau)$  while holding  $\tau$  fixed, as described above. The computation ends when  $\text{test}$  and  $|C(\mathbf{h}, \mathbf{d}) - \text{aim}| / \text{aim}$  are both smaller than 0.0001.

## Results and Discussion

Preliminary results using synthetic signals plus random noise are shown in Figures 1–3, using the constant-aim mode. Our ml1r algorithm is comparable in speed to MaxEnt, on the order of 10 ms for the spectra shown here using a 3.7 GHz Intel Xeon processor. IST is considerably slower, on the order of 100 ms for the spectra shown. The synthetic data was nonuniformly sampled using an exponentially biased scheme selecting 100 or 300 samples from the uniform 1024-element data vector. The top trace of Fig. 1 shows the DFT of the uniformly sampled data; the second row of traces depict the results of “nonuniform DFT”, in which the samples not utilized are set to zero prior to Fourier transformation. Successive traces show the results for IST, MaxEnt, and our new algorithm (ml1r, for minimized  $l_1$  norm, real), for 100 samples (left) and 300 samples (right) for IST and MaxEnt. For ml1r, because the algorithm necessarily applies one zero-fill to ensure causality of the spectrum, samples with indices larger than 512 were removed from the sampling schedule, so that the number of actual samples used for ml1r reconstructions were 83 and 250 respectively, instead of 100 and 300 samples, respectively. In Fig. 1 the MaxEnt and ml1r reconstructions were computed in the “MINT” limit, tightly constrained to match the sampled data, so that they can be compared with the IST results. While IST does much better at suppressing sampling artifacts than MaxEnt or ml1r, close examination of the broad peak reveals artifactual splitting of the single resonance for both 100 and 300 samples. ml1r attains comparable sampling artifact suppression without splitting the resonance. Fig. 2 shows results obtained with IST, MaxEnt, and ml1r in the Bayesian limit, where *aim* is set to a value comparable to the noise. Comparison with Fig. 1 shows that in the Bayesian limit minimizing the complex  $l_1$  norm (IST) does worse at suppressing sampling artifacts than in the MINT limit, whereas ml1r suppresses sampling artifacts better in the Bayesian limit; changing the value of *aim* from the MINT to the Bayesian regime has relatively little impact on MaxEnt. Fig. 3 gives a clearer indication of the impact of the choice of *aim* on the structure of the weak peak.

## Concluding remarks and future directions

Our algorithm readily extends to data of arbitrary dimension, by restricting the  $l_1$  norm computation to the purely absorptive quadrant (real-real component for two dimensions) of the spectrum, and we are conducting critical comparisons of the approach with conventional MaxEnt reconstruction and MINT for two- and three-dimensional NUS data. The preliminary one-dimensional results shown here suggest a number of avenues for further investigation. If indeed the greater sparsity of the real component is responsible for the mitigation of artifacts compared to using the  $l_1$  norm of the complex spectrum, then similar effects should be apparent when comparing reconstructions of broad spectral components with those for narrow spectral components, or comparing crowded spectra with simpler, more sparse spectra. Other issues that need further investigation are the multiplicity of solutions due to the non-strict convexity of the  $l_1$  norm, its application to spectra that contain both positive and negative signals, and the use of deconvolution to reduce the support (increase the sparsity) of NMR spectra.

While our approach facilitates critical comparison of CS and MaxEnt approaches in NMR, there remain other obstacles to identifying best practices. One is the lack of consensus metrics for spectral quality, which also hampers the comparison of NUS strategies. Further advances in NUS and non-Fourier spectrum analysis could benefit from a “shared task”, or a competition similar to Critical Assessment of Structure Prediction (CASP)(29).

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

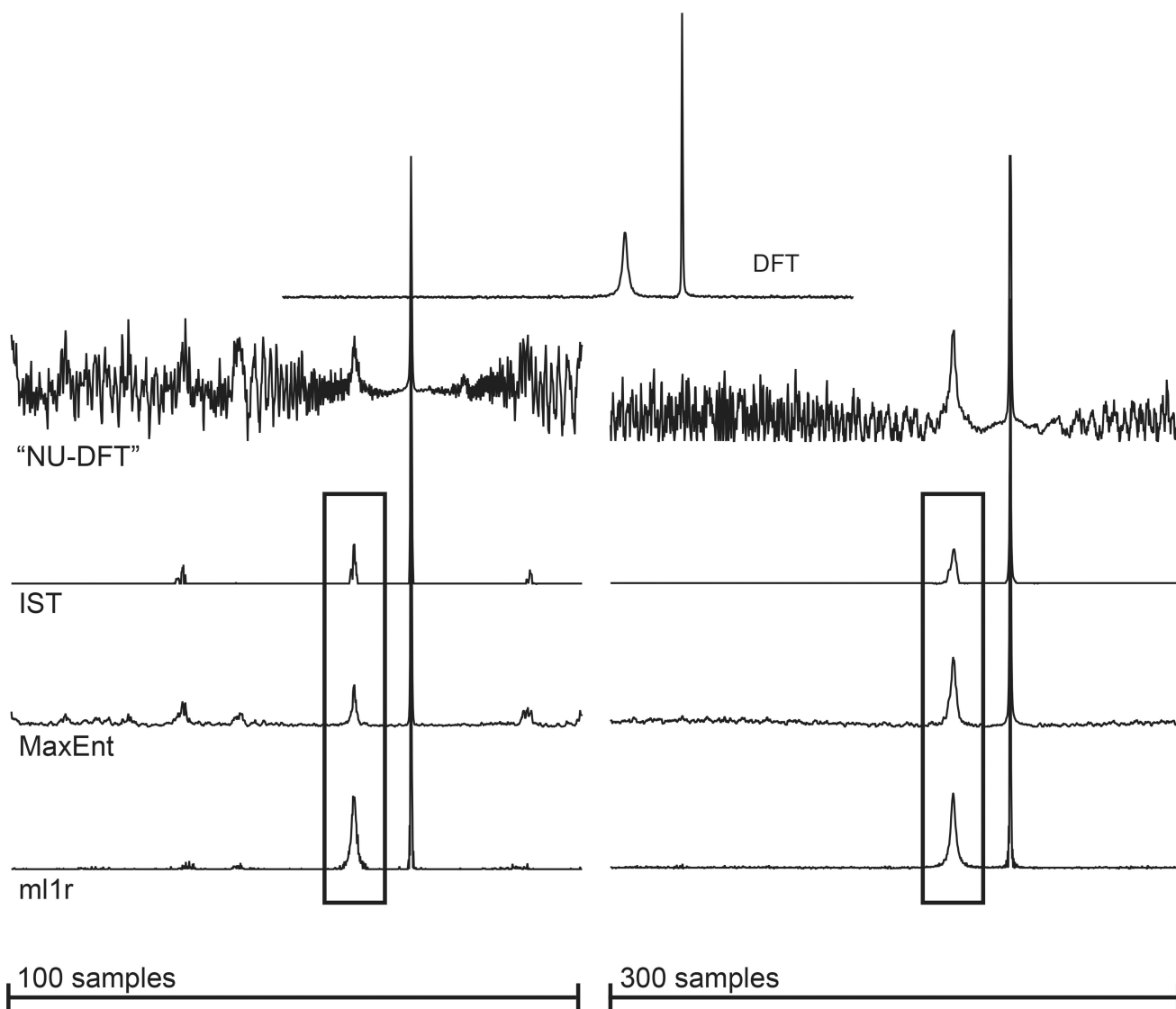
We thank David Donoho and Adam Schuyler for useful discussions. This work was supported by a grant from the US National Institutes of Health (GM104517).

## References

1. Heideman MT, Johnson DH, Burrus CS. Gauss and the History of the Fast Fourier Transform. *IEEE Acoustics, Speech, and Signal Processing*. 1984; 1:14–21.
2. Prony R. Essai expérimental et analytique, etc. *J. l'Ecole Polytechnique, Paris*. 1795; 1:24–76.
3. Fourier, JBJ. *The Analytical Theory of Heat* (translation of *Théorie Analytique de la Chaleur*). The University Press; London: 1878.
4. Cooley JW, Tukey JW. An algorithm for the machine calculation of complex Fourier series. *Math. Comput.* 1965; 19:297–301.
5. Candes EJ, Romberg J, Tao T. Robust Uncertainty Principles: Exact Signal Reconstruction From Highly Incomplete Frequency Information. *IEEE Transactions on Information Theory*. 2006; 52:489–509.
6. Mobli M, Hoch JC. Nonuniform sampling and non-Fourier signal processing methods in multidimensional NMR. *Prog Nucl Magn Reson Spectrosc.* 2014; 83C:21–41.
7. Bostock MJ, Holland DJ, Nietlispach D. Compressed sensing reconstruction of undersampled 3D NOESY spectra: application to large membrane proteins. *J Biomol NMR*. 2012; 54:15–32. [PubMed: 22833055]
8. Holland DJ, Bostock MJ, Gladden LF, Nietlispach D. Fast multidimensional NMR spectroscopy using compressed sensing. *Angew Chem Int Ed Engl.* 2011; 50:6548–6551. [PubMed: 21648028]
9. Kazimierczuk K, Orekhov VY. Accelerated NMR spectroscopy by using compressed sensing. *Angew Chem Int Ed Engl.* 2011; 50:5556–5559. [PubMed: 21538743]
10. Hyberts SG, Milbradt AG, Wagner AB, 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; 52:315–327. [PubMed: 22331404]
11. Shrot Y, Frydman L. Compressed sensing and the reconstruction of ultrafast 2D NMR data: Principles and biomolecular applications. *J Magn Reson.* 2011; 209:352–358. [PubMed: 21316276]
12. Logan, BF. Ph.D. thesis. Department of Electrical Engineering, Columbia University; 1965. Properties of high-pass signals.
13. Drori I. Fast l1 Minimization by Iterative Thresholding for Multidimensional NMR Spectroscopy. *Eurasip Journal on Advances Signal Processing*. 2007; 2007:1–10.
14. 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]
15. 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]
16. van Cittert PH. Zum Einfluss der Spaltbreite auf die Intensitätsverteilung in Spektrallinien. II. *Zeitschrift für Physik.* 1931; 69:298–308.

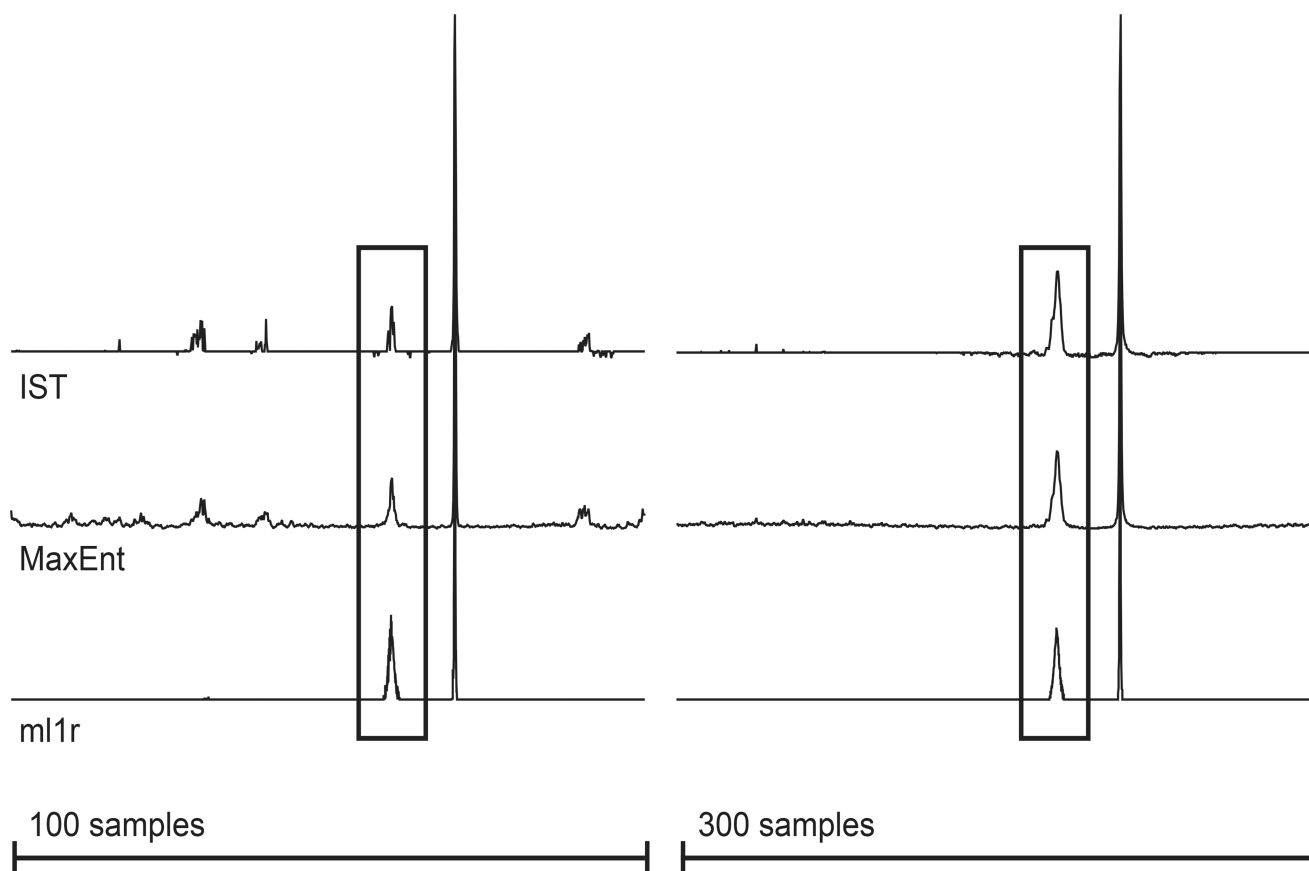


17. Jansson PA. Method for determining the response function of a high-resolution infrared spectrometer. *J. Opt. Soc. Amer.* 1970; 60:184–191.
18. Papoulis A. A new algorithm in spectral analysis and band-limited extrapolation. *IEEE Trans. Circ. Syst.* 1975; CAS-22:735–742.
19. Donoho DL, Johnstone IM, Stern AS, Hoch JC. Maximum Entropy and the Nearly Black Object (with discussion). *J Roy Stat Soc B.* 1992; 54:41–81.
20. Wernecke SJ, D' Addario LR. Maximum Entropy Image Reconstruction. *IEEE Trans. Comp.* 1977; 26:351–364.
21. Hyberts SG, et al. Ultrahigh-resolution (1)H-(13)C HSQC spectra of metabolite mixtures using nonlinear sampling and forward maximum entropy reconstruction. *J Am Chem Soc.* 2007; 129:5108–5116. [PubMed: 17388596]
22. Paramasivam S, et al. Enhanced sensitivity by nonuniform sampling enables multidimensional MAS NMR spectroscopy of protein assemblies. *J Phys Chem B.* 2012; 116:7416–7427. [PubMed: 22667827]
23. Mayzel M, Kazimierczuk K, Orekhov VY. The causality principle in the reconstruction of sparse NMR spectra. *Chem Commun (Camb).* 2014; 50:8947–8950. [PubMed: 24975496]
24. Hoch, JC., Stern, AS., et al. *NMR Data Processing.* Wiley-Liss; New York: 1996.
25. Bracewell, RN. *The Fourier Transform and Its Applications.* 3. McGraw-Hill; Singapore: 2000.
26. Marion D, Wüthrich K. Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of 1H-1H spin-spin coupling constants in proteins. *Biochem. Biophys. Res. Comm.* 1983; 113:967–974. [PubMed: 6307308]
27. Press, WH., Flannery, BP., Teukolsky, SA., Vetterling, WT. *Numerical Recipes in Fortran.* Cambridge University Press; Cambridge: 1992.
28. Schmieder P, Stern AS, Wagner G, Hoch JC. Quantification of maximum-entropy spectrum reconstructions. *J Magn Reson.* 1997; 125:332–339. [PubMed: 9144266]
29. 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.

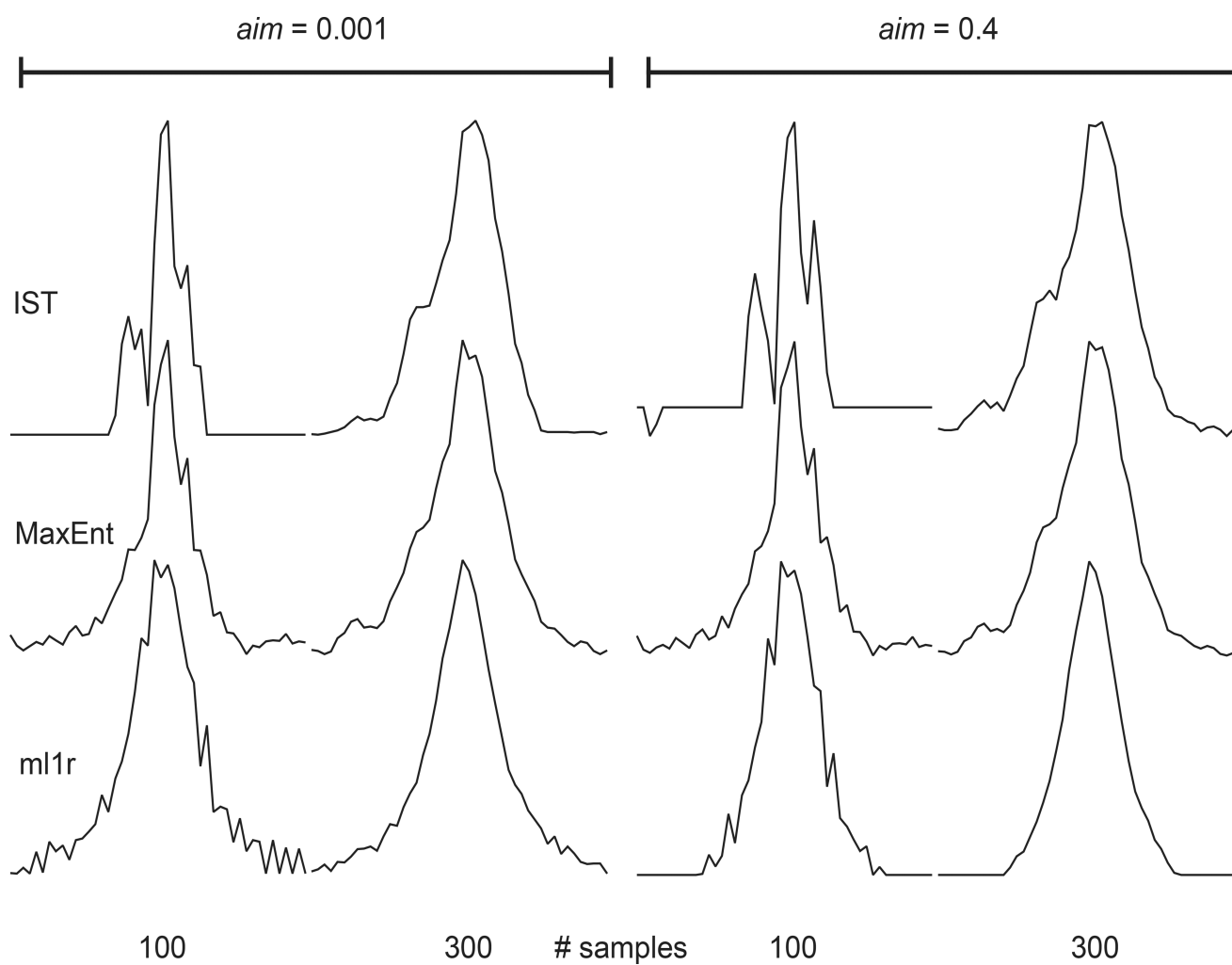


**Figure 1.**

Top row: DFT of 1024-point synthetic data plus noise (noise RMS = 0.1). Successive rows: spectral estimates employing 100 (left) and 300 (right) samples according to an exponentially biased sampling scheme (see supplementary information). The value of *aim* was  $10^{-3}$  for all IST, MaxEnt, and ml1r spectra. The peaks enclosed by rectangles are expanded in Figure 3.



**Figure 2.** Spectral estimates in the Bayesian regime employing 100 (left) and 300 (right) samples according to an exponentially biased sampling scheme (see supplementary information). The value of *aim* was 0.4 for all IST, MaxEnt, and ml1r spectra. The peaks enclosed by rectangles are expanded in Figure 3.



**Figure 3.** Expansion of the regions enclosed by rectangles in Figures 1 and 2. Spectra on the left half correspond to the MINT regime (small  $aim$ ,  $10^{-3}$ ); spectra on the right half correspond to the Bayesian regime (larger  $aim$ , 0.4).