NESTA-NMR: Efficient and Generalized Processing of Multidimensional NUS NMR Data

Michelle L. Gill, Shangjin Sun, Yifei Li, Mitchell Huang, R. Andrew Byrd
National Cancer Institute, National Institutes of Health, Frederick, MD 21702

Structural Biophysics Laboratory

Introduction

Non-uniform sampling (NUS) has been demonstrated to offer advantages over uniform sampling in the form of time savings, resolution enhancement, and, under some circumstances, sensitivity gains. Successful implementation of NUS requires two additional steps relative to uniformly sampled data: (1) selection of a sampling schedule, and (2) a method for transforming the data from the time to the frequency domain. For data transformation, ℓ1-norm (L1) regularization methods, such as iterative soft thresholding (IST) [1], have gained considerable popularity for reconstructing missing data points to create a uniformly sampled grid prior to fast Fourier transform.

We have developed a program containing a suite of regularization terms, including an L1 method that utilizes a novel gradient descent method called NESTA [2], which is shown to converge in significantly fewer steps than other IST methods and is able to reconstruct a 4D NMR spectrum in ~3 hours on desktop-class hardware. This program, called NESTA-NMR [3], is general and robust, as demonstrated by the reconstruction of both high dynamic range applications (e.g. NOESY) and a variety of J-coupled correlation experiments. NESTA-NMR also includes built-in parallelization routines, tight incorporation with NMRPipe data processing workflows, and highly customizable settings with sensible presets, providing both speed and ease of use.

Overview of Processing Steps

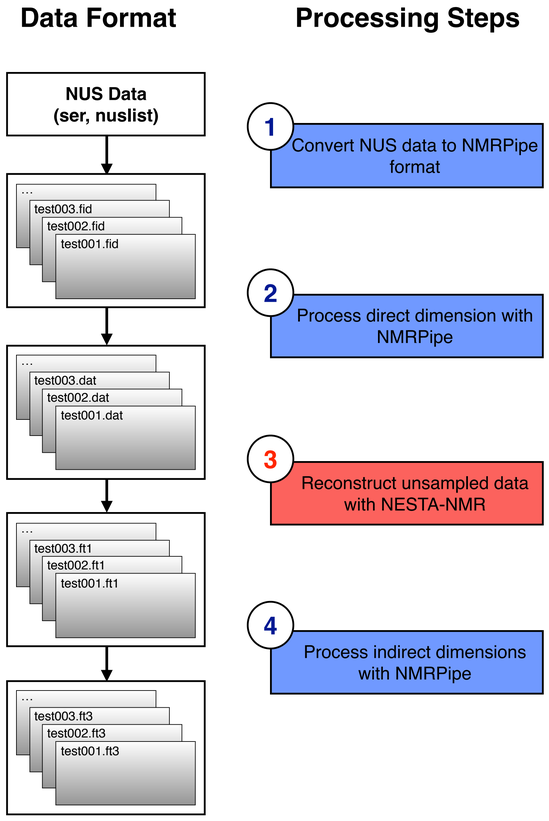


Figure 1. Processing NUS data with NMRPipe and NESTA-NMR requires only one additional step (step 3, reconstruction of unsampled data points) relative to the analogous process for uniformly sampled data. The direct dimension is transformed prior to reconstruction to simplify the computation and enable data associated with each directly detected point to be processed in an independent (parallel) fashion. Because NESTA-NMR integrates with NMRPipe, it is compatible with both Bruker and Varian/Agilent data.

References

1. Hyberts S.G., Milbradt A.G., Wagner A.B., Arthanari H. & Wagner G. (2012) *J. Biomol. NMR*, **52**, 315–27.
2. Becker S., Bobin J. & Candès E. (2011) *SIAM J. Imaging Sci.*, **4**, 1–39.
3. Sun S., Gill M., Yifei L., Huang M. & Byrd R.A. (2015) *J. Biomol. NMR*, online March 26, 2015, doi: 10.1007/s10858-015-9923-x.
4. Stern A.S., Donoho D.L. & Hoch J.C. (2007) *J. Magn. Reson.*, **188**, 295–300.
5. Drori I., (2007) *EURASIP J. Adv. Signal Process.*, 1–10.

Acknowledgments
This work was supported by the Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research.

NESTA Algorithm

Minimize $||f||_{l1}$ **subject to** $Rx = b$

L1

$$||f||_{l1} = \sum |f_k|$$

IRL1

$$||f||_{irl1} = \sum \omega_k |f_k|$$

$$|f_k| = \sqrt{f_{k,r}^2 + f_{k,i}^2}$$

$$\omega_k^{i+1} = 1/(|f_k|^i + \varepsilon)$$

Gaussian-SL0

$$||f||_{sl0} = \sum (1 - e^{-0.5|f_k|^2/\sigma^2})$$

Figure 2. The optimization performed by NESTA-NMR minimizes the L1-norm ($\|f\|_{l1}$) of the spectrum while ensuring the product of the sampled data (R) and the desired solution (x) is equal to a zero-filled matrix of the sampled data (b). NESTA-NMR contains three different regularization terms that can be minimized: L1-norm (L1), iteratively re-weighted L1-norm (IRL1), and Gaussian smoothed L0-norm (Gaussian-SL0).

Reconstruction Results

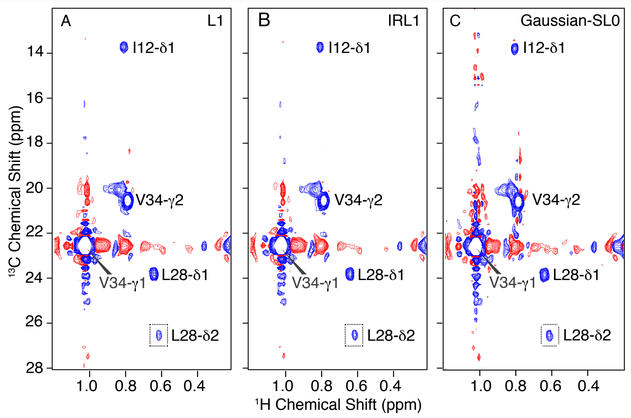
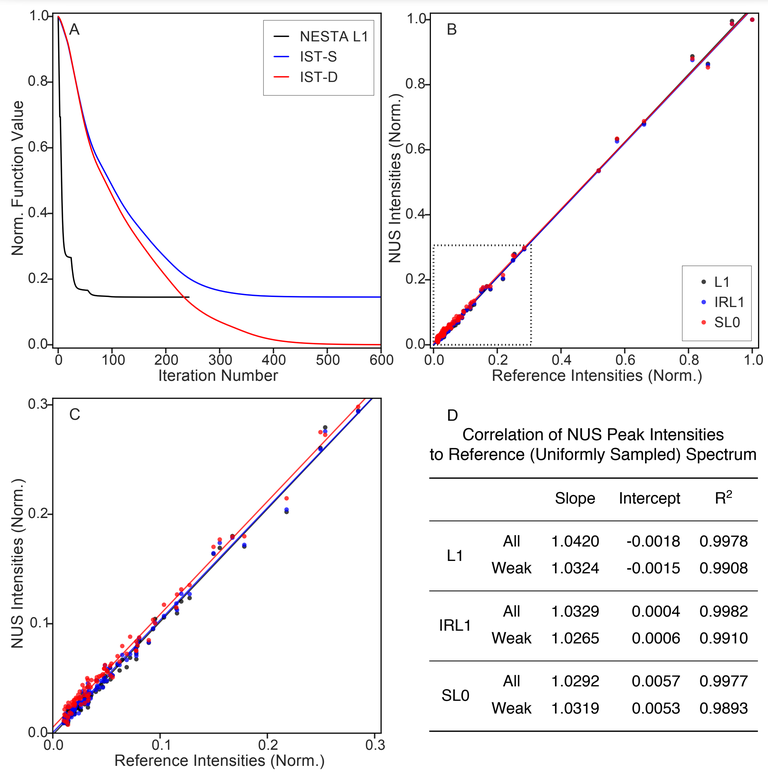


Figure 3. The reconstruction of a 4D HMQC-NOESY-HMQC spectrum demonstrates the differences between **(A)** L1, **(B)** IRL1, and **(C)** Gaussian-SL0 regularization terms in NESTA-NMR. IRL1 is slightly better than L1 at preserving weak peaks, while Gaussian-SL0 is the best, albeit with some artifacts (see also Figure 4B–D).

Convergence Rate and Accuracy



D

Correlation of NUS Peak Intensities to Reference (Uniformly Sampled) Spectrum

		Slope	Intercept	R^2
L1	All	1.0420	-0.0018	0.9978
	Weak	1.0324	-0.0015	0.9908
IRL1	All	1.0329	0.0004	0.9982
	Weak	1.0265	0.0006	0.9910
SL0	All	1.0292	0.0057	0.9977
	Weak	1.0319	0.0053	0.9893

Figure 4. NESTA [2] is a first-order gradient descent algorithm that converges in fewer iterations **(A)** than the IST methods of Stern, *et al.* (IST-S) [4] and Drori (IST-D) [5]. The difference in final function value for IST-D is related to treatment of sampled data, not reconstruction quality. **(B)** Peak intensities obtained using each of the NESTA-NMR regularization terms are consistent with those of a uniformly sampled spectrum, **(C)** including weak peaks. **(D)** The results of linear regressions on data shown in panels B and C.

Features of NESTA-NMR

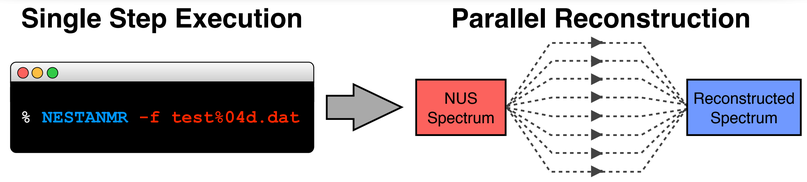


Figure 5. Execution of NESTA-NMR is performed with a single command, and reconstructions are automatically parallelized. Most reconstruction parameters are adjustable by command line flags.

Reconstruction Enhancement

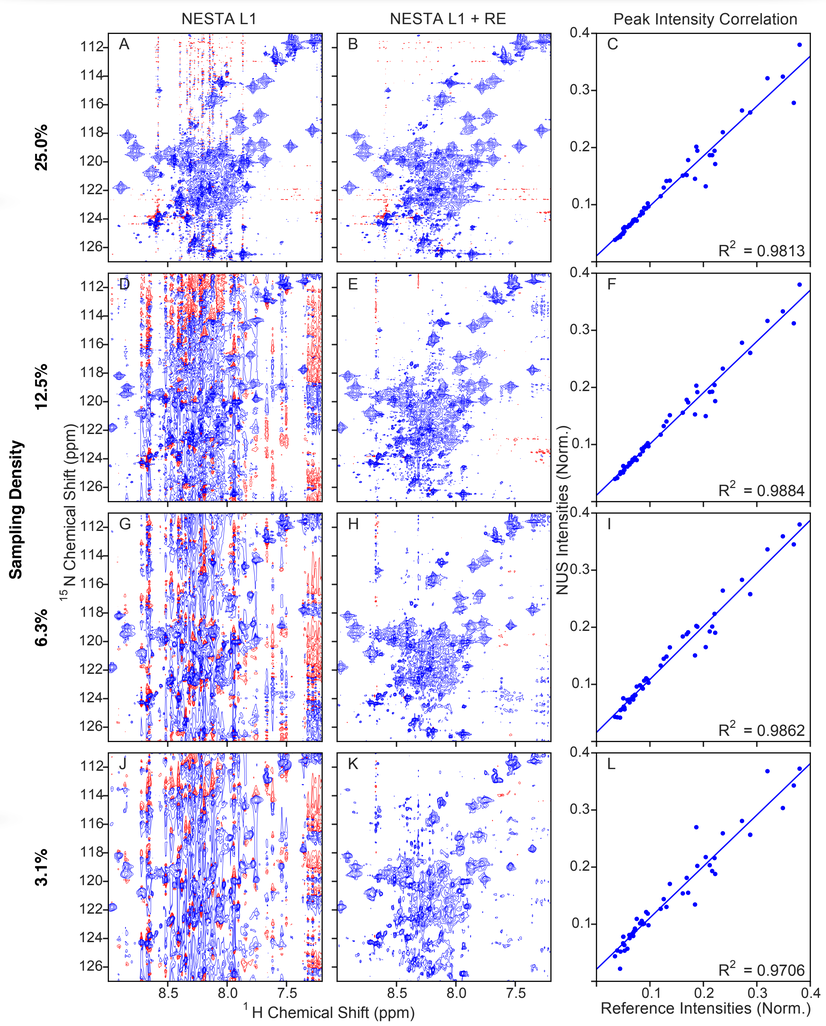


Figure 6. A forthcoming version of NESTA-NMR contains reconstruction enhancement (RE), which dramatically reduces artifacts that are problematic in cases of extreme sparsity, including lower dimensionality spectra (e.g. HSQCs and relaxation experiments). These improvements are demonstrated using a ^{1}H–^{15}N HSQC with **(A)** 25.0%, **(D)** 12.5%, **(G)** 6.3%, and **(J)** 3.1% sampling densities of an intrinsically disordered protein (IDP). IDPs often require long t_1 acquisition times to improve resolution and reduce spectral artifacts from the sharp, disordered resonances, making NUS advantageous. RE is able to remove the artifacts **(B**, **E**, **H**, and **K)**, and the resulting peak intensities demonstrate excellent correlation **(C**, **F**, **I**, and **L)** to those of a uniformly sampled reference spectrum, indicating this technique is suitable for highly quantitative methods, such as relaxation measurements.

Conclusions

- NESTA-NMR is a software suite that integrates with NMRPipe, can reconstruct NUS data in a single step, and produces very accurate results
- Incorporation of the NESTA gradient descent method enables convergence in significantly fewer iterations than other L1-norm methods
- NESTA-NMR contains three different regularization terms (L1, IRL1, and Gaussian-SL0), which can be used for reconstruction
- Execution of NESTA-NMR requires a single command and provides automatic parallelization
- Future improvements enable highly quantitative reconstruction of spectra with even lower sampling densities

NESTA-NMR can be downloaded from http://nestanmr.com