ACS OMEGA

http://pubs.acs.org/journal/acsodf

Article

New Algorithm by Maximizing Mutual Information for Correction of Frequency Drifts Arising from One-Dimensional NMR Spectroscopic Data Acquisition

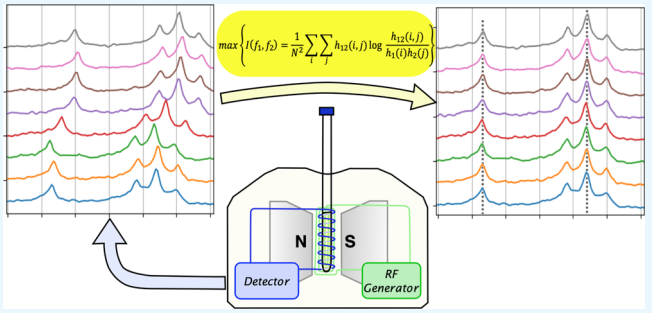
Toshi Nagata*

Cite This: *ACS Omega* 2021, 6, 31299−31304

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Benchtop nuclear magnetic resonance (NMR) instruments are getting popular these days. However, the obtained spectra sometimes suffer from significant frequency drifts, which cause difficulty in accumulating the raw data. In this paper, a new algorithm for correction of frequency drifts is proposed, which operates by maximizing mutual information between the obtained spectroscopic data. The algorithm worked well for both ^{1}H and ^{19}F NMR spectroscopic data, even in the case of very noisy ones. In comparison with the previously reported algorithms, the present algorithm has an advantage that NMR spectra complicated by signal overlapping and spin coupling can be handled without difficulty. This makes the present algorithm particularly advantageous for application of benchtop NMR spectrometers in organic chemistry.



INTRODUCTION

Benchtop nuclear magnetic resonance (NMR) spectrometers have become increasingly popular during the last decade.[1] These spectrometers utilize compact permanent magnets based on rare-earth metals[2] and have advantages of lower costs (both for installation and maintenance), smaller footprints, and safer operation (no need to use cryogens). Although the capability is limited due to the low operational frequencies, benchtop NMR spectrometers have found many important applications, including natural product analysis,[3] process control,[4] monitoring of bioreactors,[5] health sciences,[6] and education.[7]

Since these instruments are often designed for simplicity and at affordable costs, some functional components may be compromised. Among these is the field-frequency lock mechanism. When this functionality is missing, the data obtained from multiple scans cannot be accumulated directly. This spoils the advantage of Fourier-transform NMR measurements to a large degree. In particular, the educational applications of benchtop NMR spectrometers often expect that the obtained spectra can be interpreted according to the description in standard textbooks of organic chemistry such as spin−spin couplings, which requires a reasonably high resolution. In order to compensate the lack of the field-frequency lock mechanism, one needs a reasonable method for frequency drift correction between the multiple scans.

In this paper, a new algorithm for the correction of frequency drift is presented, which uses "mutual information" between the scans. Mutual information has been utilized in registering multiple two-dimensional images, and application for magnetic resonance imaging is reported.[8] The algorithm is applied for one-dimensional spectroscopic data. Several different implementations of the mutual information are presented and compared. The algorithm presented here was proven to be effective, even in the case of very noisy ^{19}F NMR spectroscopic data.

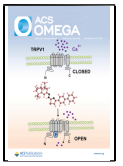
MATERIALS AND METHODS

Sample Preparation and NMR Data Acquisition. All chemicals were used as received: ethyl cinnamate and trifluoromethylbenzene (Nacalai Tesque), hexafluorobenzene (Tokyo Chemical Industries), and tetrachloromethane (Wako Pure Chemicals). Each sample solution was placed in a 5 mm Pyrex tube and set into the NMR probe without spinning.

The ^{1}H and ^{19}F NMR spectroscopic data were obtained with a Pulsar spectrometer (Oxford Instruments). The instrument utilizes a permanent magnet (1.45 T). The operating frequency was 60.3 MHz for ^{1}H and 56.8 MHz for ^{19}F. The instrument does not have a field-frequency lock mechanism.

Data acquisition was carried out with the following parameters: 90° pulse, 5.7 μs (^{1}H) and 6.3 μs (^{19}F); acquisition

Received: September 16, 2021
Accepted: October 29, 2021
Published: November 12, 2021



https://doi.org/10.1021/acsomega.1c05143
ACS Omega 2021, 6, 31299−31304

points, 16,384 (^{1}H) and 32,768 (^{19}F); and sweep width, 10 kHz (^{1}H) and 50 kHz (^{19}F). In order to compensate the initial delay of the data acquisition after the excitation pulse, FID (free induction decay) data were extrapolated before the first point by the backward linear prediction method;[9] the number of extrapolated points was eight (^{1}H) and three (^{19}F).

Algorithm for Frequency Drift Correction by Maximizing Mutual Information. The mutual information between the two spectra f_1 and f_2 is defined as follows

$$I(f_1, f_2) = \sum_{f_1} \sum_{f_2} p(f_1, f_2) \log \frac{p(f_1, f_2)}{p(f_1)p(f_2)}$$

where $p(f_1,f_2)$ is a joint probability distribution between the two spectra and $p(f_1)$ and $p(f_2)$ are marginal probability distributions of the corresponding spectra.

The probability distributions are computed from the histograms. Suppose $h_1(i)$ and $h_2(i)$ are the histograms of the spectra f_1 and f_2, respectively, and $h_{12}(i,j)$ is the joint histogram, we get

$$I(f_1, f_2) = \frac{1}{N^2} \sum_i \sum_j h_{12}(i, j) \log \frac{h_{12}(i, j)}{h_1(i)h_2(j)}$$

where N is the number of bins in calculating the histograms.

The procedure for frequency drift correction is as follows. We have a set of FID data $g_1(t)...g_M(t)$, where M is the number of scans. The spectra in the frequency domain $G_1(k)...G_M(k)$ are obtained by Fourier transform of the FID data. Now, we determine the frequency drift between the first two scans by comparing the two spectra

$$f_1 = |G_1(k)|^\gamma, \qquad f_2 = |G_2(k + \Delta k_{12})|^\gamma$$

Here, we use the absolute values of the spectra, in order to avoid the complexity of phase correction. The exponent γ (≥ 1) is an arbitrary parameter that gives more weights on larger signals and less weights on baseline noises. Finding Δk_{12} which maximizes $I(f_1,f_2)$ will give the desired frequency shift between the first two scans.

To determine the frequency shift for $g_3(t)$, we have two choices for f_1

$$\text{choice (1)}: f_1 = |G_2(k)|^\gamma, \qquad f_2 = |G_3(k + \Delta k_{23})|^\gamma$$

$$\text{choice (2)}: f_1 = |G_1(k) + G_2(k + \Delta k_{12})|^\gamma$$

$$f_2 = |G_3(k + \Delta k_{13})|^\gamma$$

Choice (1) gives the frequency shift between the second and third scans, whereas choice (2) gives the frequency shift of the third scans with respect to the "accumulated" spectrum. If the phase shift during the data acquisition is negligible, choice (2) might give a better result because the signal-to-noise ratio in f_1 is higher. We will compare the results from these two choices later.

After the frequency shifts are determined for all spectra, the corrected FIDs and spectra are computed as follows

$$g_{\text{corr},n}(t) = g_n(t) \cdot \exp\left(-\frac{2\pi t \Delta k_{1n} i}{Z}\right)$$

$$G_{\text{corr},n}(k) = G_n(k - \Delta k_{1n})$$

where Z is the number of points in each FID (and spectrum). When we use choice (1), Δk_{1n} is determined as follows

$$\Delta k_{1n} = \Delta k_{12} + ... + \Delta k_{n-1,n}$$

All calculations were performed on a MacBook computer using Python 3.9,[10] NumPy 1.21.0,[11] and Matplotlib 3.4.2.[12] The script is given in the Supporting Information.

RESULTS

Processing ^{1}H NMR Spectra with High S/N. The ^{1}H NMR spectroscopic data for ethyl cinnamate (20 mmol L^{-1} in CCl_4) were acquired with eight scans. The excerpts of the FIDs and the spectra after Fourier transform are shown in Figures 1a and 2a, respectively. The spectra in Figure 2a show significant drifts of the peak positions. We do not discuss here the cause of these drifts; instead, we will focus on how to correct these drifts in the data processing stage.

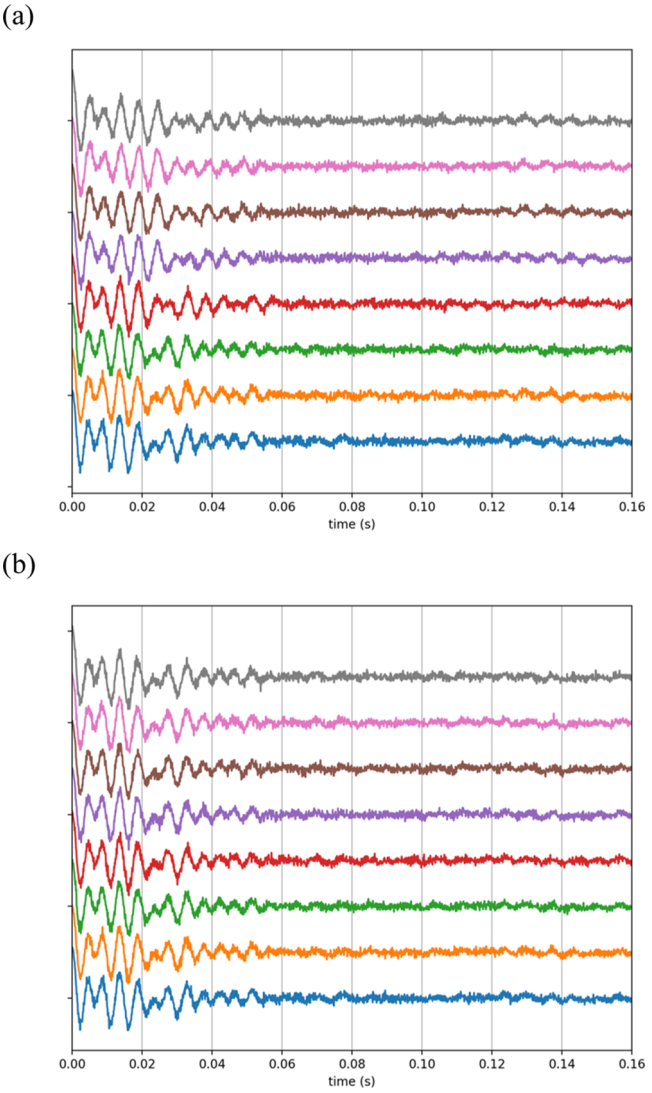


Figure 1. FIDs of the ^{1}H NMR (20 mmol L^{-1} ethyl cinnamate in CCl_4). The eight scans are stacked from bottom to top. Backward linear prediction (see the Materials and Methods) and exponential broadening (3 Hz) are applied for each FID. Only the first 1600 points (out of 16,384) are shown. (a) FIDs as obtained and (b) corrected FIDs obtained after applying the frequency offsets determined in Figure 3.

The mutual information between the two adjacent scans (i.e., choice (1) as described in the Materials and Methods section) is computed for different frequency offsets, and the results are

https://doi.org/10.1021/acsomega.1c05143
ACS Omega 2021, 6, 31299−31304

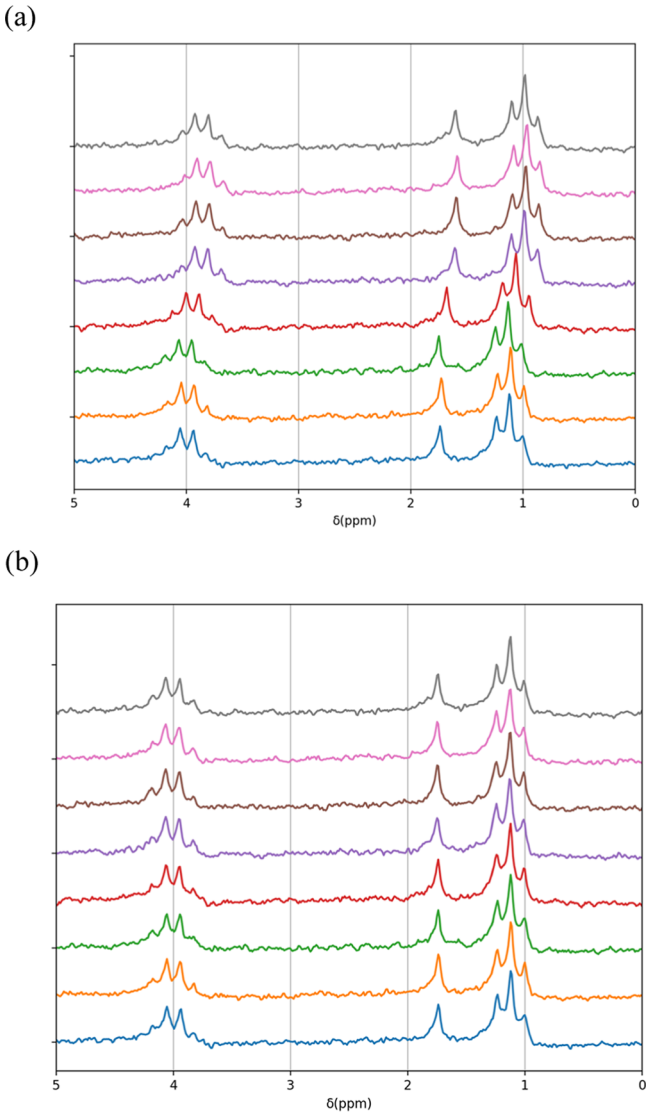


Figure 2. ^{1}H NMR spectra obtained from the FIDs in Figure 1. Only the range 0−5 ppm is shown, where the signals of the ethyl groups and contaminated water are visible. (a) FIDs are used as obtained, and (b) FIDs are used after applying the frequency offsets determined in Figure 3.

shown in Figure 3. Each trace has one maximum, which corresponds to the "best fit" between the two adjacent spectra. The maximum positions show that frequency drifts occurred at the fourth and fifth scans (corresponding to the third and fourth traces in Figure 3). These results are consistent with the spectra in Figure 2a, which also show that the fourth and fifth spectra had frequency drifts.

Using the frequency offset obtained from Figure 3, the corrected FIDs and spectra are computed. The results are shown in Figures 1b and 2b. All FIDs and spectra are in good agreement, indicating that the frequency drifts were appropriately corrected.

The frequency shifts are also determined using choice (2) as described in the Materials and Methods section (Figure 4). The results are similar to those in Figure 3, except that the traces in Figure 4 are slightly less noisy. In this experiment, both choices (1) and (2) were proven to be equally useful.

Processing ^{1}H NMR Spectra with Low S/N. The ^{1}H NMR spectroscopic data were measured for a 10 times diluted solution

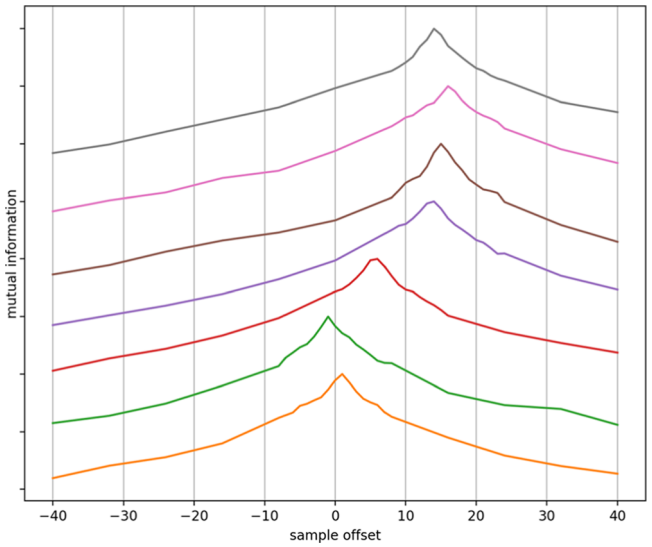


Figure 3. Mutual information based on choice (1) (i.e., between one spectrum and the preceding spectrum). The nth trace from the bottom corresponds to the mutual information between the $(n-1)$th and nth spectra. The horizontal axis denotes the frequency offset (in sampling points; 1 sample = 0.61 Hz = 0.0101 ppm) with respect to the first spectrum (rather than the offset between the two adjacent spectra), so that the offset values at the peak correspond to the spectral shift shown in Figure 2. The weighting exponent $\gamma = 1$.

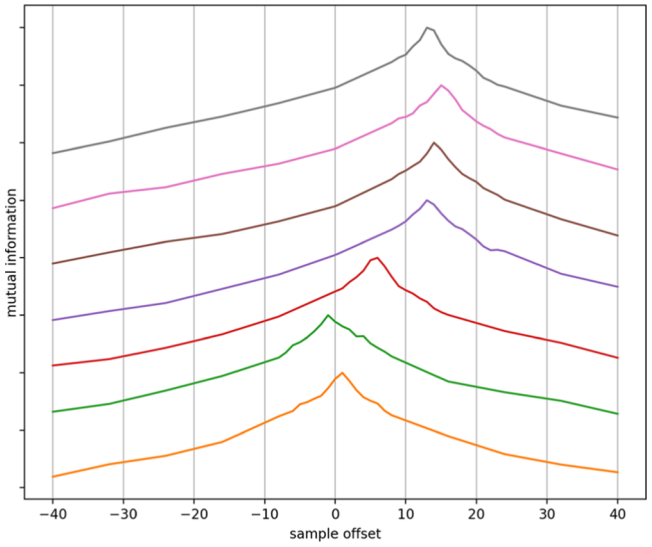


Figure 4. Mutual information based on choice (2) (i.e., between one spectrum and the corrected average of the preceding spectra). The nth trace from the bottom corresponds to the mutual information between the nth spectra and the corrected average of the first to the $(n-1)$th spectra. The weighting exponent $\gamma = 1$.

of ethyl cinnamate (2 mmol L^{-1} in CCl_4). In this case, the signal of contaminated water becomes very large in comparison with other signals. The frequency drifts were evaluated as described in the last section, and the results are shown in Figure 5a. Although the procedure mostly worked well, the corrected spectra still have a slight discrepancy in the peak-top positions. These insufficient corrections can be attributed to the "blunt" shapes of the mutual information profile, which cause ambiguity in the peak-top positions.

These "blunt" shapes of the mutual information profiles can be improved by increasing the weighting exponent γ. With a

https://doi.org/10.1021/acsomega.1c05143
ACS Omega 2021, 6, 31299−31304

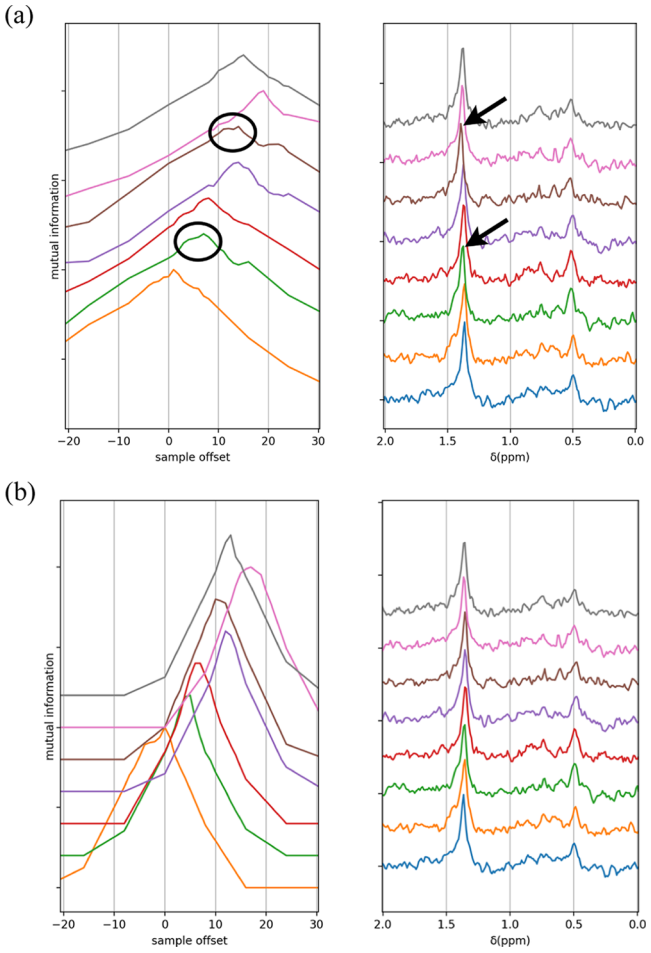


Figure 5. Processing of the ^{1}H NMR data with low S/N. Left, mutual information based on choice (2); right, the corrected spectra. (a) Results obtained with the weighting exponent $\gamma = 1$. The peaks indicated by arrows still have frequency shifts. These insufficient corrections can be attributed to the blunt shapes of the mutual information profile (marked by ovals in the left figure). (b) Results obtained with the weighting exponent $\gamma = 3$.

larger value of γ, the data points near the noise level have less weight, and the points near the strong signals have more weight. Such a weighting scheme should make the mutual information more sensitive to the frequency shift. The results with $\gamma = 3$ are shown in Figure 5b. The mutual information profiles were sharp enough to determine the frequency drift unequivocally.

Processing ^{19}F NMR Spectra. The ^{19}F NMR spectroscopic data for a mixture of trifluoromethylbenzene (20 mmol L^{-1}) and hexafluorobenzene (10 mmol L^{-1}) in CCl_4 were acquired with eight scans and processed in a similar manner as the ^{1}H NMR ones. A larger value of exponential broadening (10 Hz) was applied because the ^{19}F signals were much narrower than the ^{1}H signals due to the wider frequency range. The results are shown in Figure 6. The procedure also worked well in this case.

The ^{19}F NMR spectroscopic data of a 10 times diluted solution (trifluoromethylbenzene 2 mmol L^{-1} and hexafluorobenzene 1 mmol L^{-1} in CCl_4) were similarly acquired and processed. In this case, the spectra were noisy, and the signals were only barely visible (Figure 7). Nevertheless, detection of the frequency drifts was successful, and the spectra were appropriately corrected (Figure 8).

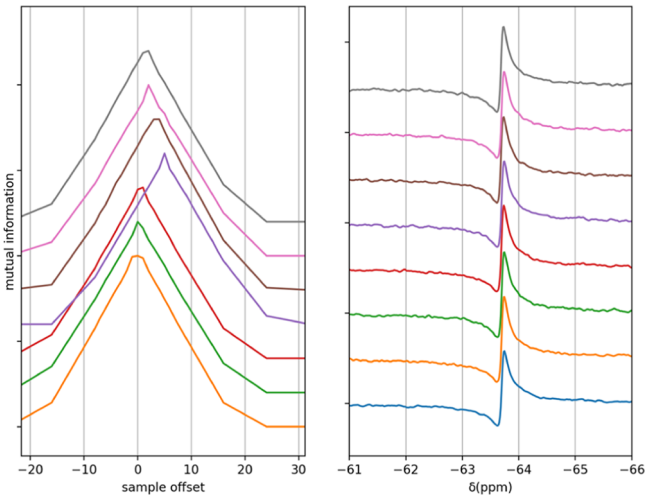


Figure 6. Processing of the ^{19}F NMR data, with the weighting exponent $\gamma = 3$. Left, mutual information based on choice (2), plotted versus the sample offset (1 sample = 1.53 Hz = 0.027 ppm). Right, the corrected spectra.

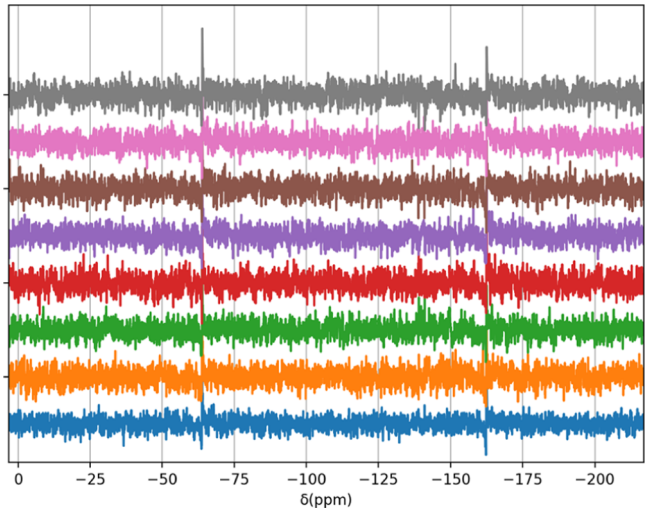


Figure 7. ^{19}F NMR spectra obtained from a dilute solution. Each trace shows a single scan.

DISCUSSION

The frequency drift is a wide-spread problem in NMR spectroscopy. It is desirable that the problem is addressed by implementation of suitable hardware.[13,14] Nevertheless, numerical correction methods are useful when the hardware solution is not applicable.

The manufacturer of our NMR spectrometer (Oxford Instruments) has developed an advanced software tool called SoftLock for correcting frequency shifts.[15] At present, we do not have access to this proprietary software, so we cannot discuss this software in comparison with our solution. We would rather compare our approach with other scientific reports that have their algorithms openly explained.

There are reports on correction algorithms based on "principal component analysis," which were successfully applied for in vivo ^{31}P NMR data.[16−18] These algorithms essentially assume that the "true" spectrum can be described as a set of appropriately shaped (e.g., Lorentzian) signals. In the case of ^{1}H NMR, however, the signals usually have complicated shapes

https://doi.org/10.1021/acsomega.1c05143
ACS Omega 2021, 6, 31299−31304

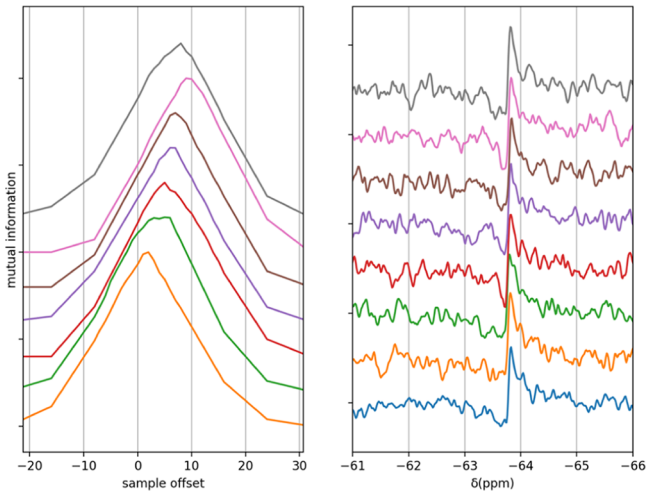


Figure 8. Processing of the ^{19}F NMR data in Figure 7, with the weighting exponent $\gamma = 3$. Left, mutual information based on choice (2); right, the corrected spectra.

because of the spin couplings. Moreover, the spectra obtained on the benchtop spectrometer often show second-order spin couplings because the dispersion of chemical shifts is not large enough in comparison with the spin coupling width.

A more recent report describes another algorithm for correction of frequency shifts, with a particular focus on low-field NMR measurements.[19] The algorithm is based on direct determination of the frequency shifts from the peak-top positions and is reported to work well for ^{1}H NMR spectra with spin coupling. However, identification of the "peak-top" positions is not trivial in complicated ^{1}H NMR spectra.

The algorithm described in this paper works well even for complicated ^{1}H NMR spectra. The concept of mutual information has been used in image recognition,[20] where the target image is treated as a whole, without assuming any structural model inside it. Measurements of ^{1}H NMR spectra, particularly in the field of organic chemistry, are more like image processing than fitting to a parametrized model function. Therefore, the algorithm presented here is particularly suitable for application in organic chemistry.

CONCLUSIONS

The mutual information algorithm is proven to be effective for correction of frequency drifts in measurements of one-dimensional NMR spectra. The algorithm worked well even in the case of very noisy spectroscopic data, in which peak detection is not trivial. This algorithm can be easily implemented as Python scripts, which can be translated into other programming languages. This processing algorithm will be of great help in utilizing benchtop NMR equipment which does not have the field-frequency lock mechanism.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.1c05143.

Python script as an implementation of the proposed new algorithm and usage of the Python script (PDF)

AUTHOR INFORMATION

Corresponding Author

Toshi Nagata − *Department of Applied Chemistry, Faculty of Science and Technology, Meijo University, Nagoya 468-8502 Aichi, Japan;* orcid.org/0000-0002-7172-9718; Email: tnagata@meijo-u.ac.jp

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.1c05143

Notes

The author declares no competing financial interest.

ACKNOWLEDGMENTS

The benchtop NMR instrument was introduced with the financial support to Meijo University from the MEXT (Ministry of Education, Culture, Sports, Science and Technology, Japan), as the Comprehensive Support Program for Innovation of Private Universities (FY2014).

REFERENCES

(1) Giberson, J.; Scicluna, J.; Legge, N.; Longstaffe, J. Developments in benchtop NMR spectroscopy 2015−2020. *Annu. Rep. NMR Spectrosc.* **2021**, *102*, 153−246.

(2) Edwards, J. C. The emergence of benchtop NMR systems and the exciting future of the technology. *Magn. Reson. Chem.* **2016**, *54*, 492−493.

(3) Beek, T. A. Low-field benchtop NMR spectroscopy: status and prospects in natural product analysis. *Phytochem. Anal.* **2021**, *32*, 24−37.

(4) Meyer, K.; Kern, S.; Zientek, N.; Guthausen, G.; Maiwald, M. Process control with compact NMR. *Trends Anal. Chem.* **2016**, *83*, 39−52.

(5) Bouillaud, D.; Drouin, D.; Charrier, B.; Jacquemmoz, C.; Farjon, J.; Giraudeau, P.; Gonçalves, O. Using benchtop NMR spectroscopy as an online non-invasive in vivo lipid sensor for microalgae cultivated in photobioreactors. *Process Biochem.* **2020**, *93*, 63−68.

(6) Percival, B. C.; Grootveld, M.; Gibson, M.; Osman, Y.; Molinari, M.; Jafari, F.; Sahota, T.; Martin, M.; Casanova, F.; Mather, M. L.; Edgar, M.; Masania, J.; Wilson, P. B. Low-Field, Benchtop NMR Spectroscopy as a Potential Tool for Point-of-Care Diagnostics of Metabolic Conditions: Validation, Protocols and Computational Models. *High-Throughput* **2018**, *8*, 2.

(7) Edgar, M.; Percival, B. C.; Gibson, M.; Masania, J.; Beresford, K.; Wilson, P. B.; Grootveld, M. Benchtop NMR Spectroscopy and Spectral Analysis of the cis- and trans-Stilbene Products of the Wittig Reaction. *J. Chem. Educ.* **2019**, *96*, 1938−1947.

(8) Positano, V.; Bernardeschi, I.; Zampa, V.; Marinelli, M.; Landini, L.; Santarelli, M. F. Automatic 2D registration of renal perfusion image sequences by mutual information and adaptive prediction. *Magn. Reson. Mater. Phys.* **2013**, *26*, 325−335.

(9) Barkhuijsen, H.; de Beer, R.; Bovée, W. M. M. J.; van Ormondt, D. Retrieval of frequencies, amplitudes, damping factors, and phases from time-domain signals using a linear least-squares procedure. *J. Magn. Reson.* **1985**, *61*, 465−481.

(10) Van Rossum, G.; Drake, F. L. *Python 3 Reference Manual*; Scotts Valley, CA: CreateSpace, 2009.

(11) Harris, C. R.; Millman, K. J.; van der Walt, S. J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N. J.; Kern, R.; Picus, M.; Hoyer, S.; van Kerkwijk, M. H.; Brett, M.; Haldane, A.; del Río, J. F.; Wiebe, M.; Peterson, P.; Gérard-Marchant, P.; Sheppard, K.; Reddy, T.; Weckesser, W.; Abbasi, H.; Gohlke, C.; Oliphant, T. E. Array programming with NumPy. *Nature* **2020**, *585*, 357−362.

(12) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 90−95.

(13) Galuppini, G.; Magni, L.; Ferrante, G. The Field-Frequency Lock for Fast Field Cycling Magnetic Resonance: From NMR to MRI. *Front. Phys.* **2021**, *9*, 688479.
(14) Chen, S.; Xu, L.; Wang, H.; Dai, S. Field-frequency lock approach for 21.3-MHz high-performance NMR relaxation analyzer. *AIP Adv.* **2018**, *8*, 075327.
(15) Oxford Instruments. Pulser, Delivering NMR to your benchtop. https://nmr.oxinst.jp/assets/uploads/products/magres/documents/Pulsar-product-information.pdf (retrieved Oct 18, 2021).
(16) Brown, T. R.; Stoyanova, R. NMR Spectral Quantitation by Principal-Component Analysis. II. Determination of Frequency and Phase Shifts. *J. Magn. Reson. B* **1996**, *112*, 32−43.
(17) Witjes, H.; Melssen, W. J.; in't Zandt, H. J. A.; van der Graaf, M.; Heerschap, A.; Buydens, L. M. C. Automatic Correction for Phase Shifts, Frequency Shifts, and Lineshape Distortions across a Series of Single Resonance Lines in Large Spectral Data Sets. *J. Magn. Reson.* **2000**, *144*, 35−44.
(18) Stoyanova, R.; Brown, T. R. NMR Spectral Quantitation by Principal Component Analysis. III. A Generalized Procedure for Determination of Lineshape Variations. *J. Magn. Reson.* **2002**, *154*, 163−175.
(19) Qui, L.; Liu, C.; Dong, H.; Xu, L.; Zhang, Y.; Krause, H.-J.; Xie, X.-M.; Offenhäusser, A. Time-Domain Frequency Correction Method for Averaging Low-Field NMR Signals Acquired in Urban Laboratory Environment. *Chin. Phys. Lett.* **2012**, *29*, 107601.
(20) Pluim, J. P. W.; Maintz, J. B. A.; Viergever, M. A. Mutual-information-based registration of medical images: a survey. *IEEE Trans. Med. Imag.* **2003**, *22*, 986−1004.

https://doi.org/10.1021/acsomega.1c05143
ACS Omega 2021, 6, 31299−31304