ACADEMIC
PRESS

Communication

An efficient algorithm for automatic phase correction of NMR spectra based on entropy minimization

Li Chen,[a] Zhiqiang Weng,[b] LaiYoong Goh,[b] and Marc Garland[c,*]

[a] *Institute of Chemical & Engineering Sciences, 28 Ayer Rajah Crescent, #02-05/08 Singapore 139959, Singapore*
[b] *Department of Chemistry, 3 Science Drive 3, National University of Singapore, Singapore 117543, Singapore*
[c] *Department of Chemical and Environmental Engineering, 4 Engineering Drive 4, National University of Singapore, Singapore 119260, Singapore*

Abstract

A new algorithm for automatic phase correction of NMR spectra based on entropy minimization is proposed. The optimal zero-order and first-order phase corrections for a NMR spectrum are determined by minimizing entropy. The objective function is constructed using a Shannon-type information entropy measure. Entropy is defined as the normalized derivative of the NMR spectral data. The algorithm has been successfully applied to experimental ^{1}H NMR spectra. The results of automatic phase correction are found to be comparable to, or perhaps better than, manual phase correction. The advantages of this automatic phase correction algorithm include its simple mathematical basis and the straightforward, reproducible, and efficient optimization procedure. The algorithm is implemented in the Matlab program ACME—Automated phase Correction based on Minimization of Entropy.
© 2002 Elsevier Science (USA). All rights reserved.

Keywords: Phase correction; NMR spectra; Entropy minimization; Derivative

1. Introduction

Normally, zero-order and first-order phase corrections are required for Fourier transform NMR spectra in order to obtain the desired appearance of the real part of the spectra. The zero-order phase misadjustment arises from the phase difference between the reference phase and the receiver detector phase. The first-order phase misadjustment arises from the time delay between excitation and detection, flip-angle variation across the spectrum, and phase shifts from the filter employed to reduce noise outside the spectral bandwidth [1–3]. Zero-order phase correction is frequency-independent, while first-order phase correction is frequency-dependent.

The misphased Fourier transform NMR spectrum consisting of the complex points can be phase corrected by applying the equations

$$R_i = R_i^0 \cos(\phi_i) - I_i^0 \sin(\phi_i), \tag{1}$$

$$I_i = I_i^0 \cos(\phi_i) + R_i^0 \sin(\phi_i), \tag{2}$$

where R_i^0 and I_i^0 represent the misphased ith data points for the real and imaginary parts of the NMR spectral data, respectively, R_i and I_i denote the phase-corrected ith data points, and ϕ_i is the total phase correction in radians applied to the ith data points as shown in

$$\phi_i = phc0 + phc1 \times \frac{i}{n}, \tag{3}$$

where $phc0$ is the zero-order term, $phc1$ is the constant part of the first-order term, and n is the total number of real data points.

There already exist algorithms automating the procedure of phase correction. As early as 1969, Ernst [4] introduced an automatic phase correction method using the Hilbert transform to find a dispersion spectrum with total integral zero and absorption spectrum with maximum integral. This approach is limited by the spectral signal-to-noise ratio and baseline distortion. Craig and Marshall [1] proposed a method based on dispersion versus absorption plots (DISPA). However, when there are no isolated spectral lines, the phases cannot be cal-

* Corresponding author. Fax: +65-6779-1936.
 E-mail address: chemvg@nus.edu.sg (M. Garland).

culated correctly because overlap distorts the DISPA circle. Van Vaals and Gerwen [5] developed a method employing the measured phases at the peak positions of spectral lines to calculate the phase distortion. Brown et al. [6] introduced an automatic method by baseline optimization. Heuer [7] reported a noniterative method, APSL (automatic phasing by symmetrizing lines), for automatic phase correction. This approach exploits the fact that symmetric lines are maximally symmetric in correctly phased absorption spectra.

Although many algorithms have been proposed for automatic phase correction, some of them involve rather complicated mathematical models and some are limited by signal-to-noise ratio or overlapping bands. In this paper, an efficient algorithm for automatic phase correction is proposed based on entropy minimization.

2. Theory

Compared with the initial Fourier transform NMR spectra, the desired phased NMR spectra have non-negative bands and show, in many ways, the simplest spectral features. However, it can be noted that the simplest representation of a signal corresponds to its entropy-minimized form (given an appropriate measure). Every probability distribution has some uncertainty associated with it. The concept of "entropy" for a quantitative measure of uncertainty was first introduced to information theory by Shannon [8]. Shannon suggested the following equation to measure the uncertainty and called it the entropy of the probability distribution h:

$$S = -\sum_j h_j \ln h_j. \tag{4}$$

Entropy minimization was first successfully applied to the estimation of pure component spectra from unknown mixture spectra by Sasaki et al. [9,10]. Entropy minimization is widely recognized as the essence of pattern recognition [11,12]:

$$\text{Min } G = -\sum_s \sum_v h_{sv} \ln h_{sv} + P(a, c). \tag{5}$$

The probability distribution h_{sv} was first defined as the normalized second derivative of the estimated and scaled pure component absorptivity, s is the number of species, v is data channels, G is the objective value, and P is a penalty function to ensure nonnegativity in the estimated pure component spectra a and concentrations c. The results showed that entropy minimization can ensure estimated pure component spectra with the smoothest and simplest features. Zeng and Garland [13] modified the approach of Sasaki et al. and applied entropy minimization to the estimation of pure component spectra from highly overlapping simulated mixture

spectra, and the expression of the entropy function with the fourth derivatives was desired in their case. Further modifications of this approach have been successfully used for system identification purposes in reactive chemical systems using FTIR spectroscopy [14–18].

As mentioned above, the derivatives of spectra are often defined to be probability distributions h in different problems of spectral analysis such as the work of Sasaki et al. [9,10] and Garland and co-workers [13–18]. Different expressions for entropy may be preferred in different problems. Therefore, it is necessary to test different expressions of entropy (1st–4th derivative) when entropy minimization is applied to the phase correction of NMR spectra.

When a Fourier transform NMR spectrum is appropriately phased, the real part of the phased spectrum has only nonnegative spectral bands and this is in many ways the simplest physically meaningful spectral representation. In stark contrast, the imaginary part possesses both positive and negative extrema [1]. Accordingly, only the entropy of real parts of phased spectra by *phc*0 and *phc*1 is considered in the objective function.

The objective function possessing a Shannon [8] type information entropy measure for the phase correction problem is shown in the following equation, where h is the normalized derivatives of the signal or NMR spectral data, P is a penalty function to ensure nonnegative bands in the spectrum, and m is the order of the derivative of the phased spectra ($m = 1$, or 2, or 3, or 4):

$$\text{Min } E = -\sum_i h_i \ln h_i + P(R_i)$$

w.r.t. *phc*0, *phc*1

$$h_i = \frac{|R_i^m|}{\sum_i |R_i^m|} \tag{6}$$

$$R_i = R_i^0 \cos(\phi_i) - I_i^0 \sin(\phi_i)$$

$$\phi_i = phc0 + phc1 \times \frac{i}{n}.$$

In the above minimization problem, the penalty function P is introduced to ensure nonnegative bands in the spectrum,

$$P(R_i) = \gamma \left[\sum_i F(R_i) R_i^2 \right], \tag{7}$$

where γ is a penalty factor which should be set appropriately to balance the contributions of the entropy and penalty parts. The function F is defined as

$$F(y) = \begin{cases} 0, & y \geqslant 0 \\ 1, & y < 0. \end{cases} \tag{8}$$

In the present contribution, the factors of the zero-order and the first-order phase correction for one Fourier transform NMR spectrum, *phc*0 and *phc*1, are optimized

by the simplex method. The simplex method is appropriate in this case because only two unknowns must be determined. More sophisticated optimization methods such as simulated annealing are totally unnecessary.

3. Experimental details and results

The experimental data used in this study arise from the reaction of $CpCr(CO)_2(S_2CNEt_2)$ and $[CpCr(CO)_3]_2$ in d_8-Toluene as solvent [19]. ^1H NMR spectra were measured on a Bruker 300 MHz FT NMR spectrometer. The digital resolution of the frequency domain spectrum was 0.726609 Hz/Pt. There are 4551 data points for each spectrum. 1D Bruker WIN-NMR software was used to analyze the data. Prior to Fourier transformation, DC correction and exponential apodization with LB = 0.1 (line broadening factor) were applied to the FID data. Then a FT ^1H NMR spectrum was obtained as shown in Fig. 1a. The automatic phase correction was done using Bruker's 1D WIN-NMR software, and the autophased spectrum is shown in Fig. 1b.

The entropy minimization algorithm for autophasing FT NMR spectra is written in MATLAB (ACME—Automated phase Correction based on Minimization of Ent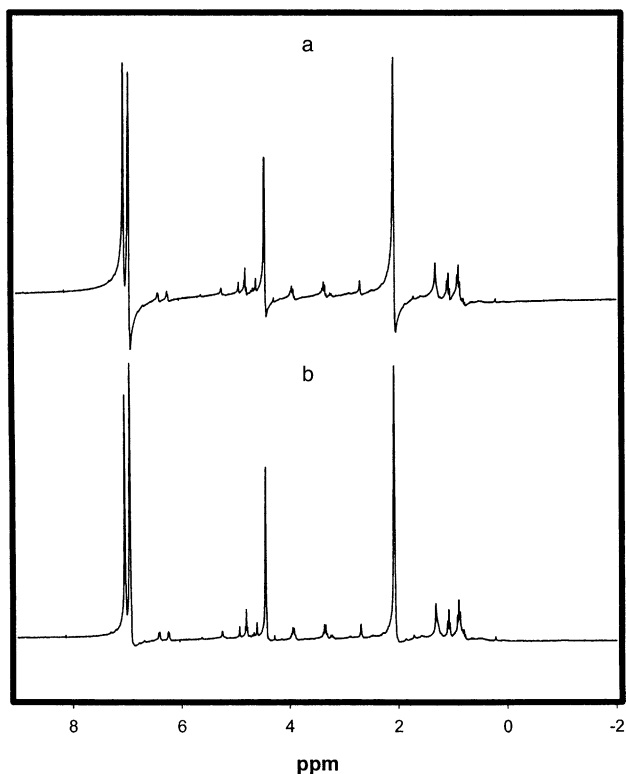ropy) [20]. The simplex method was employed to optimize the two variables *phc*0 and *phc*1 for the minimization problem. The computing time is a few seconds on a PC. In order to test the effects of noisy data on entropy minimization, randomly distributed noise was added to the spectrum in Fig. 1a. The approximate signal-to-noise ratio was 15. The noisy FT NMR spectrum is shown in Fig. 2a. Then entropy minimizations using different derivatives were applied to the FT NMR spectra in Figs. 1a and 2a. The results of the entropy minimization are shown in Table 1. The numerical values of the entropy and penalty were also calculated for the manually phased spectrum and the autophased spectrum by 1D WIN-NMR. The results are shown in Table 2.

The spectra after automatic phase correction by entropy minimization using the 1st derivative for FT NMR spectra in Figs. 1a and 2a are shown in Figs. 3 (solid curve) and 2b, respectively. In addition, the FT NMR spectrum was manually corrected using the 1D WIN-NMR software. The manually phased spectrum is also shown in Fig. 3 (dashed curve). It is noted that the

Fig. 1. A Fourier transform NMR spectrum (a) and the result (b) after automatic phase correction using WIN-NMR.
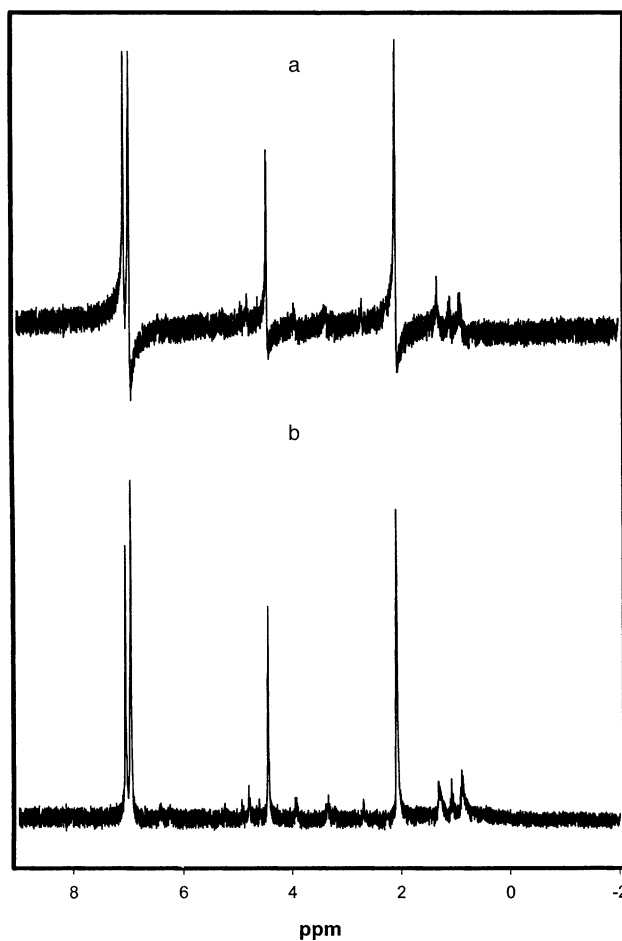

Fig. 2. A noisy Fourier transform NMR spectrum (a) and the result (b) after automatic phase correction by entropy minimization using the 1st derivative.

Table 1
The results of entropy minimization using different derivatives for FT NMR spectra in Figs. 1a and 2a

Results		1st derivative	2nd derivative	3rd derivative	4th derivative
FT NMR spectrum in Fig. 1a	Optimal *phc*0 (degree)	39.08	39.08	39.08	39.08
($\gamma = 5e - 5$ in Eq. (7))	Optimal *phc*1 (degree)	16.38	16.38	16.38	16.38
	Entropy	6.23	7.07	6.83	7.39
	Penalty	5.14	5.14	5.14	5.14
	Objective value	11.37	12.21	11.97	12.53
FT NMR spectrum in Fig. 2a	Optimal *phc*0 (degree)	38.63	38.42	38.42	38.42
($\gamma = 8e - 3$ in Eq. (7))	Optimal *phc*1 (degree)	17.86	17.84	17.84	17.80
	Entropy	8.00	8.16	8.13	8.15
	Penalty	5.53	5.42	5.42	5.42
	Objective value	13.53	13.58	13.55	13.57

Table 2
The contributions of entropy and penalty for the manually phased spectrum (Fig. 3, dashed curve) and the autophased spectrum (Fig. 1b) by 1D WIN-NMR

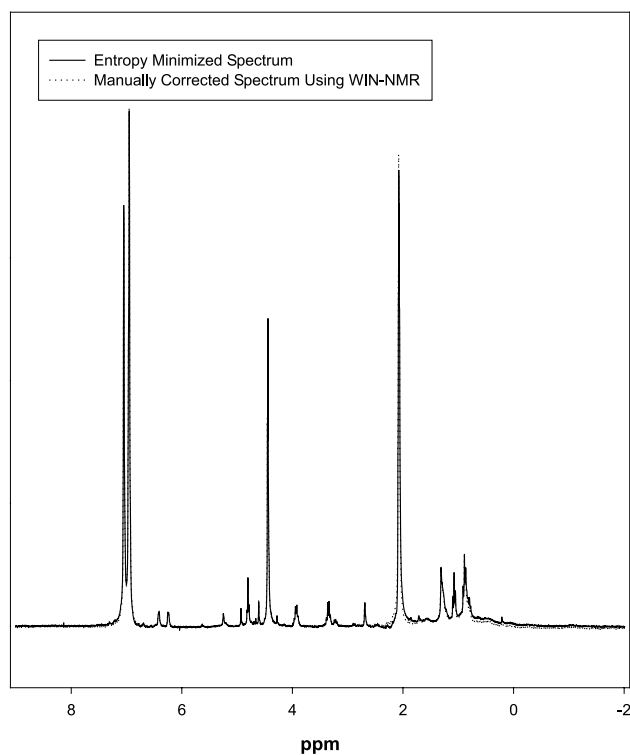| Results		1st derivative	2nd derivative	3rd derivative	4th derivative
Manually phased spectrum	Entropy	6.23	7.07	6.82	7.38
($\gamma = 5e - 5$ in Eq. (7))	Penalty	40.29	40.29	40.29	40.29
	Objective value	46.52	47.36	47.11	47.67
Autophased spectrum	Entropy	6.24	7.08	6.82	7.37
($\gamma = 5e - 5$ in Eq. (7))	Penalty	356.22	356.22	356.22	356.22
	Objective value	362.46	363.30	363.04	363.59

Fig. 3. The comparison of the result after automatic phase correction by entropy minimization using the 1st derivative (solid curve) and the result after manual phase correction using WIN-NMR (dashed curve).

autophased spectrum using entropy minimization is much better than the one using default automatic phase correction in 1D WIN-NMR (Fig. 1b) and is very similar to the manually phased spectrum (dashed curve in Fig. 3).

As shown in Table 1, it is noted that all derivatives offer similar results for phase correction even in the case of very noisy data. All derivatives work very well for the phase correction problems. In addition, as shown in Table 2, it is found that the entropy of the manually phased spectrum is close to that of the autophased spectrum achieved by entropy minimization, but the penalty of the former is much larger than that of the latter, which means that there are more negative data in the manually phased spectrum. Meanwhile, the autophased spectrum by 1D WIN-NMR is also compared with the above discussed spectra in terms of the contributions of entropy and penalty. Similarly, it is noted that the penalty of the autophased spectrum by 1D WIN-NMR is much higher than that of the autophased spectrum by entropy minimization. There are many negative data in the autophased spectrum by 1D WIN-NMR.

4. Further discussion

The present algorithm is, in and of itself, an interesting alternative to previous methods for phase correction in NMR spectroscopy. However, it is interesting to note

that the motivation for this development actually arose from the need to process hundred, even thousands, of sequential NMR spectra obtained from reaction studies. The resulting NMR spectra, having the important characteristic of minimized entropy, are thus properly preconditioned for further numerical analysis—particularly system identification of reactive systems [21], and pattern recognition/deconvolution using entropy measures [11–17].

5. Conclusions

Automatic phase correction for FT NMR spectra drew considerable attention in the 1990s [22–26]. Various automatic algorithms for phase correction had been proposed using DISPA, symmetrizing lines, baseline optimization, etc. These algorithms are limited by signal-to-noise ratio, overlapping bands, etc. In contrast, the automatic phase correction algorithm presented in this contribution only employs the rather straightforward concept of signal entropy. Entropy minimization can efficiently generate the desired autophased spectra even in the presence of considerable noise. The expressions for the entropy function with 1st–4th derivatives all offer desired results for phase correction of FT NMR spectra. The algorithm as implemented in MatLab is fast.

References

[1] E.C. Craig, A.G. Marshall, J. Magn. Reson. 76 (1988) 458.

[2] B.L. Neff, J.L. Ackerman, J.S. Waugh, J. Magn. Reson. 25 (1977) 335.

[3] J.-M. Daubenfeld, J.-C. Boubel, J.-J. Delpuech, B. Neff, J.C. Escalier, J. Magn. Reson. 62 (1985) 195.

[4] R.R. Ernst, J. Magn. Reson. 1 (1969) 7.

[5] J.J. van Vaals, P.H.J. van Gerwen, J. Magn. Reson. 86 (1990) 127.

[6] D.E. Brown, T.W. Campbell, R.N. Moore, J. Magn. Reson. 85 (1989) 15.

[7] A. Heuer, J. Magn. Reson. 91 (1991) 241–253.

[8] C.E. Shannon, Bell Syst. Tech. J. 3 (1948) 379–423.

[9] K. Sasaki, S. Kawata, S. Minami, Appl. Opt. 22 (1983) 3599.

[10] K. Sasaki, S. Kawata, S. Minami, Appl. Opt. 23 (1984) 1955.

[11] J.N. Kapur, Maximum-Entropy Models in Science and Engineering, Wiley Eastern, New Delhi, 1989.

[12] S. Watanabe, Pattern Recognition 13 (1981) 381.

[13] Y.Z. Zeng, M. Garland, Anal. Chim. Acta 359 (1998) 303.

[14] Y. Pan, L. Susithra, M. Garland, J. Chemometrics 14 (2000) 63.

[15] E. Widjaja, M. Garland, J. Comput. Chem. 23 (2002) 911.

[16] E. Widjaja, C. Li, M. Garland, Organometallics 21 (2002) 1991.

[17] W. Chew, E. Widjaja, M. Garland, Organometallics 21 (2002) 1982.

[18] (a) L. Chen, M. Garland, Appl. Spectrosc., in press (2002).
(b) L. Chen, Ph.D. thesis, National University of Singapore, (2002).

[19] L.Y. Goh, Z. Weng, W.K. Leong, P.H. Leung, Angew. Chem. Int. Ed. 40 (2001) 3226.

[20] The Matlab source code can be downloaded. Available from <http://www.chee.nus.edu.sg/staff/garland.html>.

[21] M. Garland, E. Visser, P. Terwiesch, D.W.T. Rippin, Anal. Chim. Acta 351 (1997) 337.

[22] R.E. Hoffman, F. Delaglio, G.C. Levy, J. Magn. Reson. 98 (1992) 231.

[23] N. Miyabayashi, Bunseki Kagaku 44 (1995) 549.

[24] G. Balacco, J. Magn. Reson. 110 (1994) 19.

[25] C. Chen, L. Kan, Comput. Methods Programs Biomed. 26 (1988) 81.

[26] H.T. Hu, Q.N. Van, V.A. Mandelshtam, A.J. Shaka, J. Magn. Reson. 134 (1998) 76.