Journal of Biomolecular NMR, 6 (1995) 361–365
ESCOM

J-Bio NMR 312

Use of the Cadzow procedure in 2D NMR for the reduction of t_1 noise

Caroline Brissac, Thérèse E. Malliavin and Marc A. Delsuc*

Centre de Biochimie Structurale, INSERM U414-CNRS UMR9955, Faculté de Pharmacie, Université Montpellier-I, 15 Avenue Ch. Flahault, F-34060 Montpellier, France

Received 7 July 1995
Accepted 5 September 1995

Keywords: t_1 Noise; 2D NMR spectroscopy; Cadzow procedure; Singular value decomposition; Gradient

Summary

A data processing approach is proposed for reducing the t_1 noise observed in multidimensional NMR spectra. This method is based on the use of the Cadzow procedure [Cadzow, J.A. (1988) *IEEE Trans. Acous. Speech Signal Proc.*, **36**, 49–62], and is demonstrated to be efficient for simulated cases as well as real experiments.

Introduction

Instabilities of the NMR spectrometer during the acquisition of a high-resolution multidimensional experiment lead to well-known artefacts, commonly denoted as 't_1 noise'. These artefacts are nearly unavoidable in long-term experiments, and can completely obscure the spectra. They are particularly awkward in the case of difference spectroscopy, such as natural abundance HMQC (Bax et al., 1983) or HMBC (Bax and Summers, 1986), or in the case of high dynamic experiments.

In this paper, we present a processing approach that permits the removal of t_1 noise. The proposed technique is based on the Cadzow procedure (Cadzow, 1988), and proves to be efficient for simulated cases as well as real experiments.

Theory

The Cadzow procedure is based on singular value decomposition of a Toeplitz matrix built from the 1D NMR time signal (the FID). This is similar to the first step of an HSVD analysis (Barkhuijsen et al., 1987). This matrix has the property that, in the case of a signal composed of N exponentially damped sinusoids with additive white noise, its N first singular values are related to the signal, whereas the remaining non-null singular values are related to the additive noise. Zeroing the noise-related singular values has the effect of reducing the noise, while leaving the signal untouched. However, truncating the singular value series to N leads to a matrix that does not retain the initial Toeplitz structure, therefore some regularisation of this matrix has to be applied to restore the initial structure. A FID can then be reconstructed from this matrix, which retains the signal information, but has lost some of the noise features. The procedure as proposed by Cadzow consists in iterating over this singular value truncation and performing regularisation steps until some convergence is reached.

This procedure is actually equivalent to a parametric modelling of the signal as a sum of N exponentially damped sinusoids, which can be efficiently performed in the time domain. The deviations of the data from this model are smoothed out by the procedure, while the signal is kept essentially untouched.

Similar principles have already been used in high-resolution NMR, in order to remove unwanted signals. The singular value decomposition of the 2D spectrum matrix (Brown and Campbell, 1990) was shown to remove artefacts by zeroing the largest singular values; the noise can be reduced by zeroing the smallest singular values. Another method, formally different from singular value decomposition but built from similar principles, was used to remove the water signal in a 3D spectrum (Mitshang et al., 1991).

The Cadzow procedure has been recently introduced into NMR processing as a means for reducing the noise present in ^{31}P in vivo spectra, thus permitting a more

*To whom correspondence should be addressed.

efficient quantisation of the observed metabolites (Diop et al., 1992). It has been shown (Diop et al., 1994) that the procedure, followed by a complete LP-SVD analysis, permits an estimation of the signal parameters without bias, and that the noise is filtered out by this operation in an optimal fashion. However, the number of lines present in the spectrum has to be estimated in order to apply the procedure. This is usually possible for in vivo spectroscopy, but might turn out to be a problem in the case of high-resolution multidimensional NMR as considered here. We present below a simple procedure, providing a rough estimate of the number of lines in the signal, which permitted us to apply the Cadzow procedure to remove the t_1 noise in a high-resolution NMR frame.

As often mentioned by many authors, the phenomenon called t_1 noise (Nagayama et al., 1978; Mehlkopf et al., 1984) is not a regular additive noise. It cannot be modelled by adding a random process to the signal of interest. Rather, it can be seen as a perturbation of the acquired signal due to instrumental instabilities during the acquisition of the experiment. In the typical case where this perturbation is one of the global parameters of the system, such as pulse phases, sample temperature, etc., the t_1 noise appears as characteristic modulations, correlated between the different signals of the spectrum (Gibbs et al., 1993). This noise is a scintillation noise (Marshall and Verdun, 1990), and has the very specific feature of being proportional to the perturbed signal. When this kind of noise occurs, acquiring more data does not improve the signal-to-noise ratio, or may even worsen it.

One way of modelling this t_1 noise process is to consider it as a random perturbation of, for instance, the intensity of the signal measured in the t_1 dimension; i.e., as a 'multiplicative' noise of the t_1 FIDs. This rough model permits us to take into account some of the characteristics of the t_1 noise, such as correlation between signals, non-cancellation by averaging, etc. It also provides a basis to simulate data polluted with t_1 noise.

In this paper, the Cadzow procedure is first applied in simulated cases, to check whether the method permits to compensate for t_1 noise as efficiently as in the case of regular noise. Then, a procedure that permits the application of this method to real 2D experiments is presented and tested.

Materials and Methods

In the present work, the Cadzow procedure was implemented as described above. A Toeplitz matrix was constructed from the given FID, using a prediction order of 100. The singular value decomposition of the resulting matrix was computed by using the ZSVDC procedure from the LINPACK package (Dongarra et al., 1979). The singular values were modified in the following way. The noise present in the FID was evaluated from the rejected singular values by calculating their quadratic mean value. The conserved N first singular values were then corrected by subtracting from them the square root of this noise value. Finally, the other singular values were set to 0. The whole procedure was implemented using commands of the LP-SVD module of the GIFA NMR-processing program (Delsuc, 1989).

To test the procedure, a simulated signal consisting of nine lines was constructed as described in Table 1 and Fig. 1. Simulated t_1 noise was added to this signal by using a multiplicative noise, as described above. However, a strictly lessening noise was used, as we could not imagine a t_1 noise source that would not decrease the signal present in a given t_1 FID. In other words, we added simulated t_1 noise to the initial noise-free signal X_i in the following way:

$$X_i' = X_i\,(1 - |\alpha\, n_i|)$$

TABLE 1
LINE PARAMETER VALUES OF SIMULATED SPECTRA AND FROM MONTE CARLO ANALYSES WITH AND WITHOUT THE CADZOW PROCEDURE

	(A) Reference[a]			(B) Line fitting[b]			(C) Cadzow[c]		
	amp	freq	lw	<amp>	<freq>	<lw>	<amp>	<freq>	<lw>
α	820.28	497.67	6.33	820.3 ± 15.8	486.30 ± 0.16	6.35 ± 0.51	820.3 ± 13.2	485.82 ± 0.14	6.38 ± 0.21
β	775.13	251.83	6.25	775.2 ± 71.5	258.86 ± 0.22	6.36 ± 1.33	775.1 ± 16.2	247.67 ± 0.17	6.48 ± 0.61
γ	415.72	347.06	6.31	415.7 ± 14.2	340.37 ± 0.18	6.35 ± 0.58	415.7 ± 12.4	340.00 ± 0.22	6.33 ± 0.31
δ	378.58	756.17	6.31	378.6 ± 30.5	739.75 ± 0.05	6.35 ± 0.51	378.6 ± 22.1	738.48 ± 0.04	6.37 ± 0.26
ε	357.47	596.76	6.36	357.4 ± 12.7	584.67 ± 0.13	6.39 ± 0.38	357.5 ± 12.0	584.89 ± 0.11	6.38 ± 0.28
η	325.37	993.53	6.36	325.4 ± 14.5	972.88 ± 0.06	6.36 ± 0.22	325.4 ± 14.5	972.94 ± 0.06	6.38 ± 0.12
χ	292.24	655.55	6.32	292.2 ± 39.3	647.30 ± 0.05	6.36 ± 0.74	292.2 ± 19.2	642.45 ± 0.05	6.37 ± 0.35
μ	246.07	299.70	6.31	246.1 ± 34.3	306.04 ± 0.24	6.32 ± 1.06	246.1 ± 17.5	295.34 ± 0.18	6.40 ± 0.56
ν	213.93	857.36	6.31	213.9 ± 32.5	840.26 ± 0.04	6.35 ± 0.49	213.9 ± 20.4	838.34 ± 0.04	6.34 ± 0.25

amp = amplitude (a.u.); freq = frequency (Hz); lw = line width (Hz).
[a] Line parameter values used to simulate the spectra.
[b] Line parameter values (mean and standard deviation) from a Monte Carlo analysis of the line-fitting procedure, performed on the noisy data sets.
[c] Line parameter values (mean and standard deviation) from a Monte Carlo analysis of the line-fitting procedure, performed on the noisy data sets after applying the Cadzow procedure.

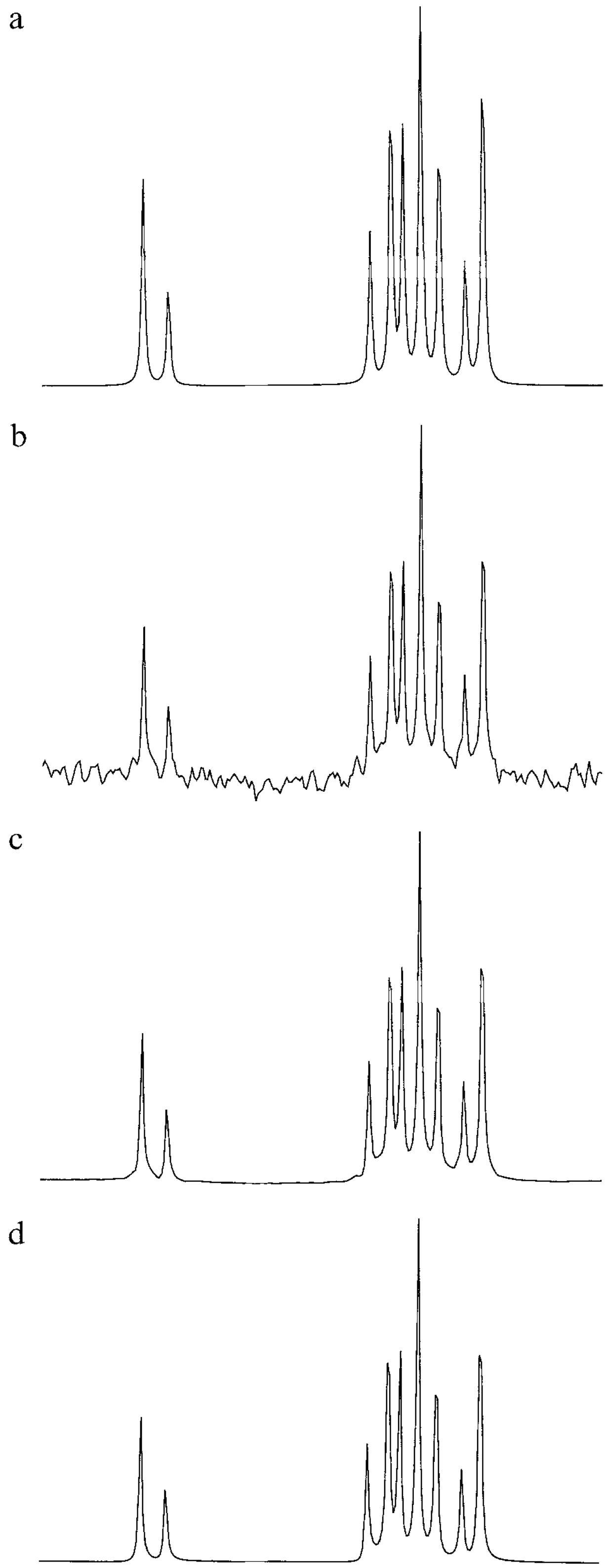

Fig. 1. The Cadzow procedure, tested on a simulated spectrum composed of nine lines. The line parameters are given in part A of Table 1. (a) The Fourier transformed noise-free simulated signal. The spectrum is simulated with a spectral width of 1000 Hz. (b) The Fourier transform of the same signal as in (a) after introduction of t_1 noise, simulated as described in the text. The noise level α is 0.25. (c) The result of the first Cadzow iteration on the data. (d) The result of five Cadzow iterations on the data.

where α is a positive constant, n_i is a random variable with a Gaussian distribution and zero mean, and $|a|$ denotes the absolute value of a. This model can probably be contested; however, we think that it is probably a more stringent test than the centred multiplicative test. On the other hand, it is probably hopeless to try to simulate real t_1 noise in a dependable manner.

This t_1 noise simulation procedure effectively reduced the total amount of signal present in the FID. Thus, to permit comparisons with the initial values, the FID was rescaled in order to keep the sum of the square of the data points constant. Figure 1 presents the result of the Cadzow procedure for $\alpha = 0.25$ and N (number of lines to find) equal to 9.

In order to evaluate the efficiency of the Cadzow procedure, the resulting spectra have been fitted to a sum of Lorentzian lines using a Levenberg–Marquart least-squares algorithm (Press et al., 1986).

To explore the characteristics of the method, we performed a Monte Carlo study on this simulated signal by using 100 realizations of the simulated noise with $\alpha = 0.15$. Each realization was then processed with two iterations of the Cadzow procedure, as described above. Finally, the resulting FID was Fourier transformed and the spectrum was fitted as a sum of nine Lorentzian lines. The mean and the standard deviation of the parameters thus obtained were then compared to the real values used for the simulation. The results are shown in Table 1.

The procedure was tested on 2D spectra by sequentially applying it to each t_1 FID, after Fourier transformation of the data set in the t_2 dimension. An estimation of the number of lines contained in each t_1 FID was obtained in the following way. The t_1 FID was temporarily Fourier transformed and phased in order to obtain a spectrum. The standard deviation of the noise was evaluated in a line-free region. A standard peak-picker was then used to detect all local maxima higher than β times the noise level. The number of peaks thus detected was used as input for the Cadzow procedure. Thus, the only user-defined parameter was β, the threshold for the peak-picker. When no peak was found by the peak-picker, no further processing was performed and the column was left unmodified in the 2D matrix.

The procedure was applied on a ^{13}C HMQC spectrum of a 10 mM sample of parvalbumin at natural abundance. The β factor value was chosen to be 4; this value was determined by trial and error.

The complete processing of the 2D data set as described above took 1 h and 51 min on an HP735/99 UNIX workstation.

Results and Discussion

The results of the Cadzow procedure obtained on a simulated noisy data set are shown in Fig. 1. It appears

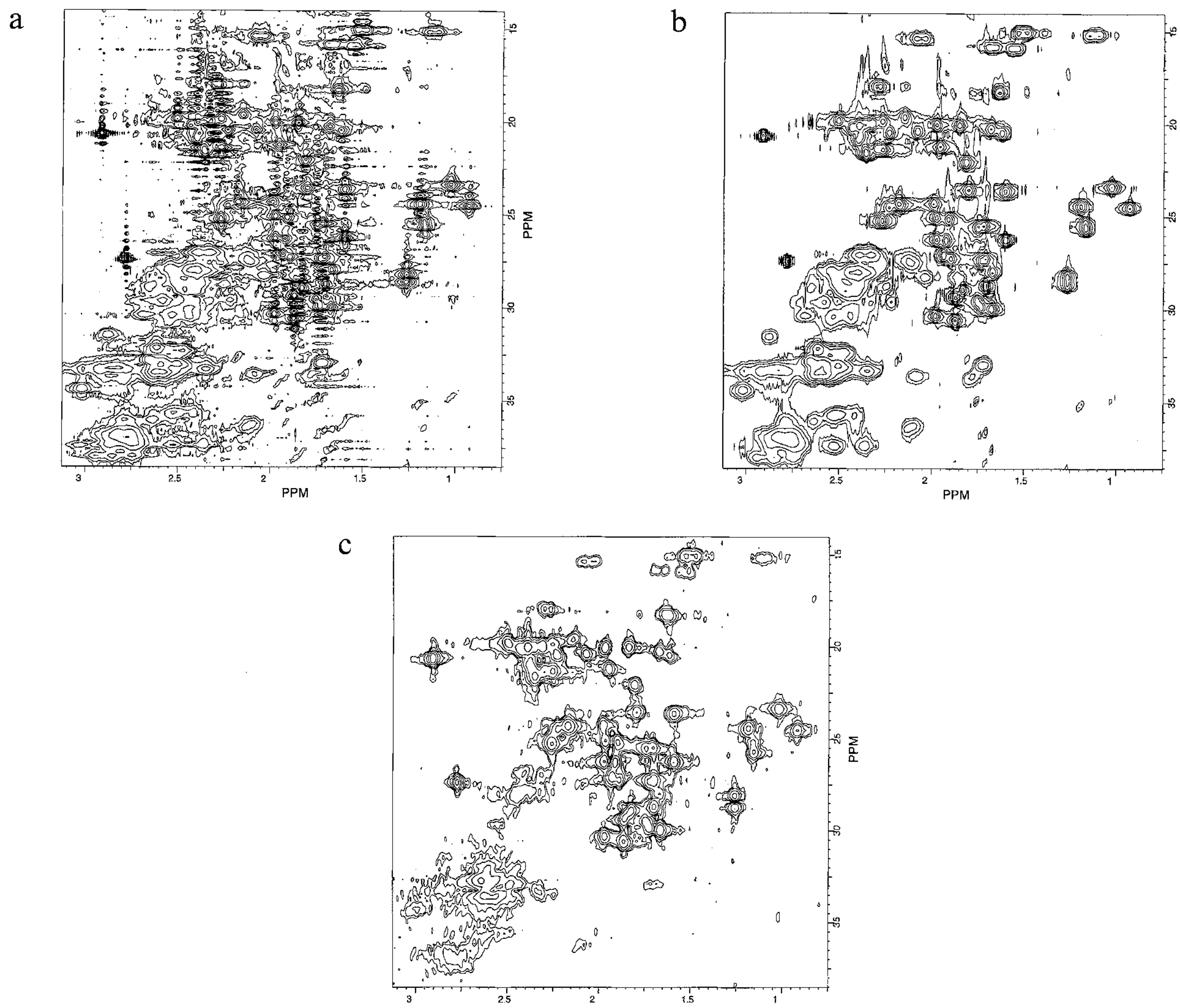

Fig. 2. Example of the use of the Cadzow procedure on a 2D natural abundance ^{13}C HMQC spectrum of pike parvalbumin. The spectrum is the result of 512 experiments, with 2048 points in the F2 direction; 16 scans for each experiment were collected, which resulted in a total measuring time of 5 h and 14 min on a Bruker AMX 400. (a) The aliphatic region, as obtained with regular processing. The experiment was recorded with the States–Haberkorn scheme. (b) The same region after application of the Cadzow procedure. (c) The same region, obtained with a PFG-HMQC experiment. The experiment was recorded with phase modulation.

from this simulation that the procedure performs successfully with simulated multiplicative noise. It can be seen in the figure that, even after the first iteration, the noise is already considerably reduced, and that subsequent iterations still further reduce it.

The Monte Carlo procedure was used to evaluate the degree of imprecision, and eventually of bias, introduced by the method (Table 1). The Cadzow procedure showed to be unbiased, as can be seen from the mean value for each line parameter extracted after the processing (Table 1, part C). The spread in the fitted parameters obtained after the Cadzow procedure tends to be much smaller than that obtained if no Cadzow procedure is applied (Table 1, parts B and C). This is an indication that some noise is indeed removed by the operation.

The simulation study shows that the Cadzow procedure improves the quality of NMR signals corrupted by t_1 noise. This result is similar to earlier observations made on processing regular additive noise (Brown and Campbell, 1990; Diop et al., 1994). The total absence of spurious signals during processing, and the very high accuracy in the frequencies of the smoothed lines should also be pointed out.

In the case of 2D spectra, a procedure based on a simple peak detection was applied to estimate the number of lines in the spectrum. The peak detection approach can be used in this case because, in a typical 2D NMR high-resolution data set, only a small number of lines is expected in each t_1 FID after the first Fourier transformation in F2.

Figure 2 shows the effect of the 2D Cadzow procedure on a natural abundance ^{13}C HSQC spectrum of parvalbumin. This kind of spectrum gives rise to large t_1 noise effects, because of imperfect cancellation of the remaining signals related to ^{12}C-attached protons (Fig. 2a). In this figure, t_1 noise can be seen as large vertical perturbations. Figure 2b demonstrates that most of the spurious signals due to t_1 noise have been correctly removed. It can also be seen that no real peak has been reduced or modified by the procedure. Moreover, it should be noted that many weak peaks, buried under the t_1 noise, are detected and that the global shape of the peaks is restored. A pulsed-field gradient experiment (Fig. 2c), naturally exempt of t_1 noise, has been performed on the sample to verify that all the peaks found in the Cadzow procedure were indeed real, and that no major peak was lost in the process. Some differences appear between Figs. 2b and c. These probably result from the application of different apodisation functions on both spectra, due to the different acquisition protocols used.

The ability of the technique to actively isolate peaks that are buried in strong t_1 noise contrasts with previously reported techniques (Manoleras and Norton, 1992), which were based on zeroing the most corrupted regions.

Recently Gibbs et al. (1993) presented a method permitting reduction of t_1 noise. Their approach is based on the correlation that appears on the noise of different signals in the spectrum. Indeed, when the t_1 noise source is the random variation of a global parameter, such as the phase of a pulse, one expects a correlated perturbation of the different t_1 noises. However, if the main perturbation is, for instance, the temperature instability, this correlation will be related to the temperature dependence of each line, and thus will be difficult to correct.

Conclusions

The noise rejection technique known as the Cadzow procedure has been applied to t_1 noise-corrupted NMR data. Tests on a simulated data set showed that the approach works, even though there is no theoretical proof for it. The procedure was then successfully applied to a real 2D HMQC experiment. The use of a pulsed-field gradient version of the same experiment verified that the improvement obtained by this technique was correct, i.e., only real peaks were found and no major peaks were lost.

This t_1 noise correction technique, based on the Cadzow procedure, advantageously compares with previously proposed techniques when considering its ability to extract real signals from intensely corrupted regions, and its independence relative to the actual form of the t_1 noise perturbation. The procedure is defined by only one user-defined parameter, which depends on the intensity of the t_1 noise.

Acknowledgements

We thank Adrien Cavé for providing the sample of pike parvalbumin. CNRS and INSERM are acknowledged for funding.

References

Barkhuijsen, H., De Beer, R. and Van Ormondt, D. (1987) *J. Magn. Reson.*, **73**, 553–557.

Bax, A., Griffey, R.H. and Hawkins, B.L. (1983) *J. Magn. Reson.*, **55**, 301–315.

Bax, A. and Summers, M.F. (1986) *J. Am. Chem. Soc.*, **108**, 2093–2094.

Brown, D.E. and Campbell, T.W. (1990) *J. Magn. Reson.*, **89**, 255–264.

Cadzow, J.A. (1988) *IEEE Trans. Acous. Speech Signal Proc.*, **36**, 49–62.

Delsuc, M.A. (1989) In *Maximum Entropy and Bayesian Methods* (Ed., Skilling, J.), Kluwer, Dordrecht, pp. 285–290.

Diop, A., Briguet, A. and Graveron-Demilly, D. (1992) *Magn. Reson. Med.*, **27**, 318–328.

Diop, A., Zaimwadghiri, Y., Briguet, A. and Graveron-Demilly, D. (1994) *J. Magn. Reson. Ser. B*, **105**, 17–24.

Dongarra, J.J., Bunch, J.R., Moler, C.B. and Stewart, G.W. (1979) *LINPACK User's Guide*, SIAM, Philadelphia, PA.

Gibbs, A., Morris, G.A., Swanson, A.G. and Cowburn, D. (1993) *J. Magn. Reson. Ser. A*, **101**, 351–356.

Manoleras, N. and Norton, R.S. (1992) *J. Biomol. NMR*, **2**, 485–494.

Marshall, A.G. and Verdun, F.R. (1990) *Fourier Transform in NMR, Optical, and Mass Spectroscopy. A User's Handbook*, Elsevier, Amsterdam.

Mehlkopf, A.F., Korbee, D. and Tiggelman, T.A. (1984) J. Magn. Reson., **58**, 315–323.

Mitshang, L., Cieslar, C., Holak, T.A. and Oschkinat, H. (1991) *J. Magn. Reson.*, **92**, 208–217.

Nagayama, K., Bachman, P., Wüthrich, K. and Ernst, R.R. (1978) *J. Magn. Reson.*, **31**, 133–148.

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1986) *Numerical Recipes*, Cambridge University Press, New York, NY.