Received: 8 November 2017 | Revised: 31 January 2018 | Accepted: 3 February 2018

DOI: 10.1002/mrc.4723

RESEARCH ARTICLE

WILEY

A subband Steiglitz-McBride algorithm for automatic analysis of FID data

M.A.R. Anjum | Pawel A. Dmochowski | Paul D. Teal

School of Engineering and Computer Science, Victoria University of Wellington, Wellington 6140, New Zealand

Correspondence
M.A.R. Anjum, School of Engineering and Computer Science, Victoria University of Wellington, Wellington 6140, New Zealand.
Email: ali@ecs.vuw.ac.nz

Fast, accurate, and automatic extraction of parameters of nuclear magnetic resonance free induction decay (FID) signal for chemical spectroscopy is a challenging problem. Recently, the Steiglitz-McBride algorithm has been shown to exhibit superior performance in terms of speed, accuracy, and automation when applied to the extraction of T_2 relaxation parameters for myelin water imaging of brain. Applying it to FID data reveals that it falls short of the second objective, the accuracy. Especially, it struggles with the issue of missed spectral peaks when applied to chemical samples with relatively dense frequency spectra. To overcome this issue, a preprocessing stage of subband decomposition is proposed before the application of Steiglitz-McBride algorithm to the FID signal. It is demonstrated that by doing so, a considerable improvement in accuracy is achieved. But this is not gained at the cost of the first objective, the speed. An adaptive subband decomposition is employed in conjunction with the Bayesian information criteria to carry out an efficient decomposition according to spectral content of the signal under investigation. Furthermore, adaptive subband decomposition and the Bayesian information criteria also serve to make the resulting algorithm independent of user input, which also fulfills the third objective, the automation. This makes the proposed algorithm favorable for fast, accurate, and automatic extraction of FID signal parameters.

KEYWORDS
adaptive, algorithm, automatic, BIC, decomposition, FID, IPF, NMR, Steiglitz-McBride, subband

1 | INTRODUCTION

One of the classic problems in magnetic resonance spectroscopy is the analysis of the free induction decay (FID) signal. In parametric approaches, the FID is modeled as a sum of complex, damped sinusoids (cisoids) buried in complex additive white Gaussian noise.[1] The general objective is to estimate the number as well as the parameters of the cisoids in the FID signal, that is, complex amplitudes, damping factors, and frequencies.

In practice, however, the task suffers from a number of problems. First, the number of cisoids in the signal model is often unknown. Second, the number of samples is quite large. Third, magnetic field inhomogeneities and variations in the chemical environment increase the complexity of the generated FID signal. All these problems make the accurate extraction of FID parameters a challenging task.[2]

Furthermore, this problem has also applications in physics, chemistry, biology, and engineering.[3]

1.1 | Prior work

Many techniques are available in the literature that attempt to address FID analysis, such as maximum

entropy, linear prediction, and state-space based methods.[4–6] Broadly, these methods can be categorized in two groups: nonparametric and parametric.

Most popular of the nonparametric methods is the Fourier transform. This has been the most obvious approach for the estimation of FID parameters for a long time. But the drawbacks of this approach are equally well established, including limited resolution and large estimate variance.[7] These problems are further exacerbated when, due to magnetic inhomogeneities and low magnetic field strength, the spectra of the cisoids overlap strongly. All these problems suggest the need for an alternative method that can cope with these issues in a robust manner.

Parametric time-domain methods, due to their superior resolution, have been shown to surpass their nonparametric counterparts.[8] Several parametric estimators for multiple-damped exponentials—such as filter diagonalization method,[9] Hankel singular value decomposition,[10] Hankel total least squares,[6] complete reduction to amplitude frequency table (CRAFT),[11] and matrix pencil[12]—have been shown to accurately quantitate and extract the parameters of the FID signal.[13] However, many of these methods (though not CRAFT) require the number of cisoids in the signal model to be known a priori, and this can introduce a systematic bias into their estimates. Furthermore, due to the use of SVD, these methods can be computationally expensive for large data lengths. However, a localized version of matrix pencil method has recently been proposed, which promises to overcome this issue.[14]

Recently, the Steiglitz-McBride method (SM) has been applied to extract the parameters of an exponential signal model.[15] This method has been shown to exhibit superior performance over most of the existing methods when applied to the extraction of T_2 relaxation parameters for myelin water imaging of the brain—a problem that employs such an exponential signal model.[8] Also, it has been shown to be completely user-input free. These advantages make the SM method an excellent choice for the extraction of FID parameters.

However, as we will show, the SM method, when applied to the FID signal, suffers from notable performance issues, especially the issue of missed spectral peaks. This can be mainly attributed to three causes. First, the signal-to-noise ratio (SNR) of the FID signal can be low if the magnetic field strength is low. Second, the spectral resolution of the FID signal can be limited due to mutual overlap of resonance peaks. Third, the length of the FID signal can be quite large because the time taken for the signal to fade into the noise can be quite long.[16] As a result, the SM method struggles for performance when applied to the FID signal.

1.2 | Contributions

In order to rectify these problems, we suggest a preprocessing stage of subband decomposition before the application of the SM method to the FID signal in order to recover its parameters. We show that this proposition solves the above three issues because the subband decomposition results in an increase in the subband SNR, an increase in spectral resolution, and a decrease in the length of the subband signals,[17] which makes the execution of the SM method in individual subbands much faster. This serves to offset any of the additional computational overheads that may be incurred due to the subband decomposition process. Finally, we employ an adaptive subband decomposition process[18] and incorporate the Bayesian information criteria (BIC) for model order selection[19] in order to make the algorithm completely independent of user input.

2 | PROBLEM FORMULATION

2.1 | Signal model

The FID signal is modeled as a sum of N-damped complex sinusoids (cisoids), being observed as M samples, regularly spaced in time by Δ:

$$r(m) = \sum_{n=1}^{N} \zeta_n e^{(-\alpha_n + j\omega_n)m\Delta} + v(m), \tag{1}$$

where $m=0,...,M-1$, $j = \sqrt{-1}$ and $\zeta_n = x_n + jy_n$ is the complex amplitude of the nth component. $\alpha_n > 0$ and ω_n are its decay rate and frequency, respectively. The noise process $v(m)$ is assumed white and complex normal, having zero mean and variance σ^2. We assume that sampling rate $1/\Delta$ is sufficient for all frequency components to be adequately represented in the sampled signal. Also, we assume that M is sufficiently high for the signal to have faded by the Mth sample, that is, $M\Delta\alpha_n \gg 1 \forall n$.

2.2 | Least square formulation

The problem is to estimate $\{\zeta_n, \alpha_n, \omega_n, N\}$. This estimation problem is generally modeled as the least-square (LS) fitting procedure[20]:

$$\min_{\{\zeta_n, p_n\}_{n=1}^{N}} \sum_{m=0}^{M-1} \left| r(m) - \sum_{n=1}^{N} \zeta_n p_n^m \right|^2 \tag{2}$$

with $p_n = e^{(-\alpha_n + j\omega_n)\Delta}$. However, this minimization problem is nonlinear and highly ill conditioned, and N is unknown.

2.3 | ARMA representation

Under the assumption that $v(m)$ is white, 1 can be modeled as an autoregressive moving average (ARMA) process of the form[21]:

$$r(m) = -\sum_{i=1}^{N} a(i)r(m-i) + \sum_{i=0}^{N-1} b(i)v(m-i), \tag{3}$$

which has been called the minimal ARMA representation of $r(m)$. However, the information of the noiseless part of 1 remains in the AR part of 2.

3 | STEIGLITZ-MCBRIDE ALGORITHM

The Steiglitz-McBride algorithm (SMA), also known as the iterative prefiltering algorithm, was originally introduced for computing LS estimates of linear system parameters. It does so by polynomial reparameterization of the problem 2 using the ARMA representation in 3.[22]

$$\min_{\{a(i)\}_{i=1}^{N},\{b(i)\}_{i=0}^{N-1}} \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|R(z)-\frac{B(z))}{A(z))}\right|^2 d\omega, \tag{4}$$

where $B(z)$ and $A(z)$ are represented by

$$B(z) = b(0) + b(1)z^{-1} + \cdots + b(N-1)z^{-N+1} \tag{5}$$

$$A(z) = 1 + a(1)z^{-1} + \cdots + a(N)z^{-N} \tag{6}$$

such that

$$\frac{B(z)}{A(z)} = \sum_{n=1}^{N}\frac{\zeta_n}{1-p_n z^{-1}}. \tag{7}$$

The SMA recursively computes the estimate of $A(z)$ by solving the following equation:

$$\min_{\{a(i)\}_{i=1}^{N},\{b(i)\}_{i=0}^{N-1}} \sum_{m=0}^{M-1}\left(\frac{R(z)A(z)-B(z)}{\hat{A}(z)}\right)^2, \tag{8}$$

where $A(z)$ is the updated estimate and $\hat{A}(z)$ is the initial estimate. $\hat{A}(z)$ is initialized to 1. This process is repeated until no further improvement in the residue is observed. Then the roots of $\hat{A}(z)$ are computed to obtain $\hat{p}_n$. Finally, $\hat{\zeta}_n$ are computed using

$$\hat{\zeta}_n = \left[(1-\hat{p}_n z^{-1})\frac{B(z)}{\hat{A}(z)}\right]_{z=\hat{p}_n}. \tag{9}$$

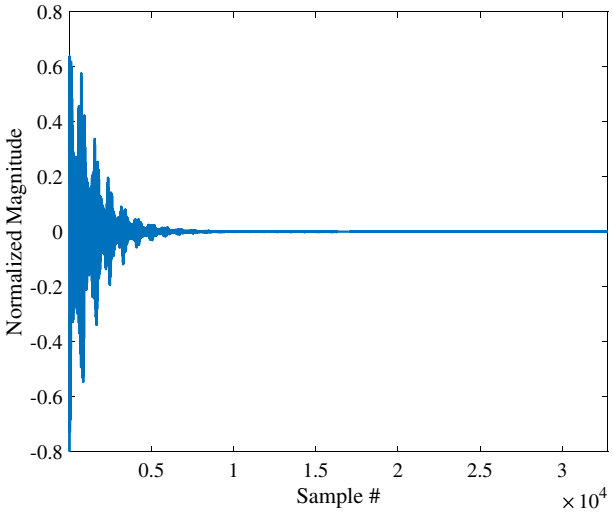

FIGURE 2 Free induction decay signal for ethanol (real part)

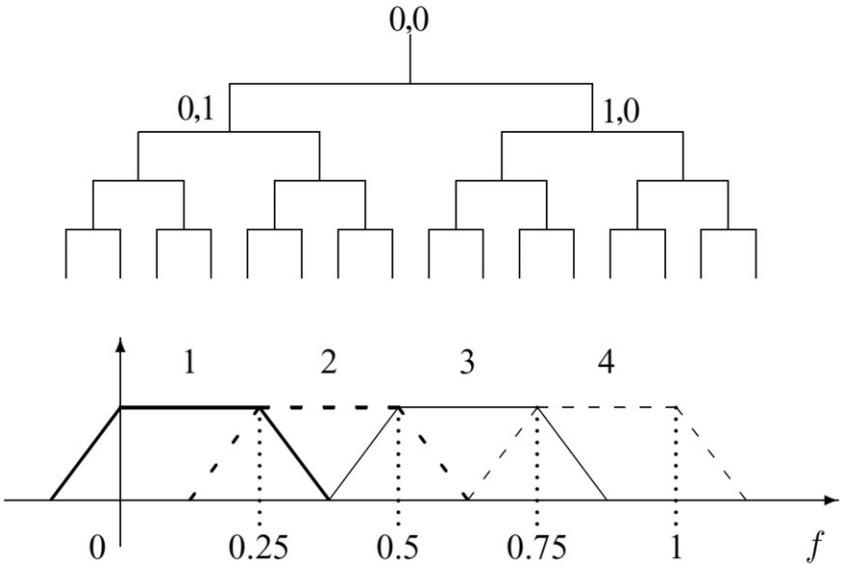

FIGURE 1 Illustration of the employed filterbank structure and the decomposition stages[18]

4 | MODEL ORDER SELECTION

SMA requires model order input since the true value of N in 1 is generally unknown. For this purpose, the BIC is considered due to its superior performance[19] when compared with alternatives such as Akaike information criterion and generalized information criterion. BIC adds an extra term to the LS cost function in 2.

$$M\log\left[\sum_{m=0}^{M-1}|r(m)-\sum_{n=1}^{N}\zeta_n p_n^m|^2\right]+2N\log(M). \quad (10)$$

The extra term penalizes higher order models if they do not yield a significant improvement in the residue. This prevents spurious estimates caused by noise, which do not actually contribute to the reconstruction of $r(m)$. Furthermore, this extra term does not incur a significant computational cost.

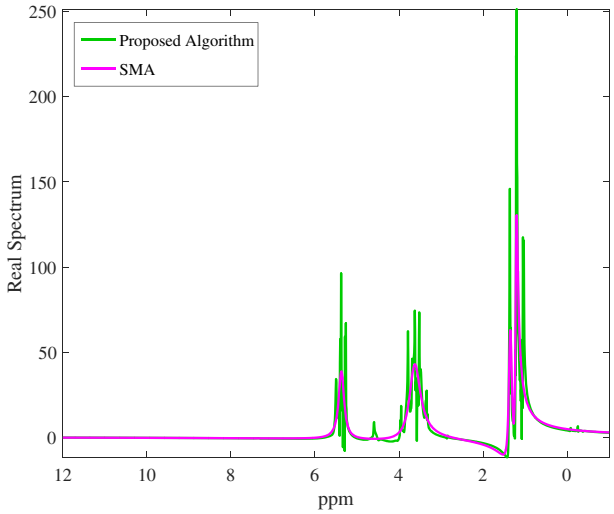

FIGURE 3 Comparison of regenerated free induction decay spectra of ethanol to the original one. At this scale, no difference is visible between the true ethanol spectrum and that reconstructed using the proposed algorithm. SMA = Steiglitz-McBride algorithm

5 | ADAPTIVE SUBBAND DECOMPOSITION

Subband decomposition reduces the fullband problem to several subband problems of lower model order. This provides advantages of an increase in SNR, an increase in frequency resolution, a decrease in signal length, and a decrease in computational complexity.[17] This gives an edge to the estimating algorithm in terms of speed and accuracy compared with the fullband problem. We adopt a uniform, multistep decomposition, in which the signal is successively filtered and decimated by a factor of 2 at each stage. Illustration of filterbank structure and the

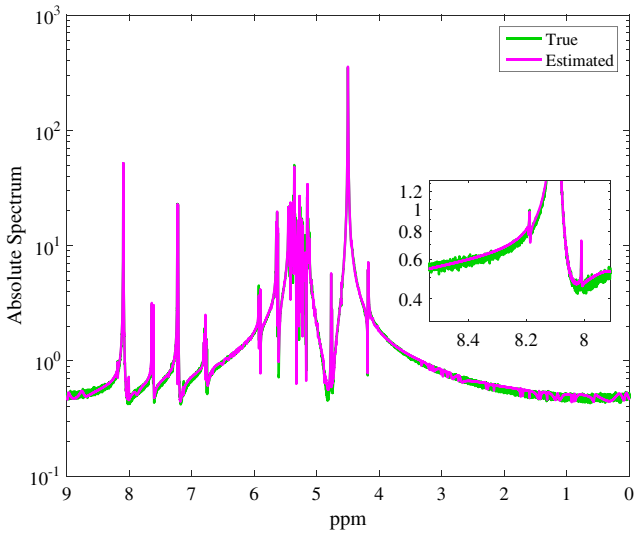

FIGURE 4 Comparison of the original and estimated ^{1}H spectrum of the lactose signal. Inset displays the zoomed view of the smallest peak captured by the proposed approach. This peak is almost 35,000 times smaller than the largest peak in the signal. Proposed algorithm took just 5.7 s to capture all 149 peaks in the signal. Mean squared error achieved was 2.5×10^{-5}

TABLE 1 A comparison of experimental results

Data		Proposed algorithm					SMA			LCModel		
Chemical sample	Expected peaks	Time (s)	MSE (10^{-6})	Depth	Missed peaks	Excess peaks	Time (s)	MSE (10^{-6})	Missed peaks	Time (s)	MSE (10^{-6})	Missed peaks
Alcohol	7	2.7	5.7	9	0	3	1.2	820	2	1.8	6.3	1
Methanol	5	1.8	0.4	8	0	2	0.3	124	3	0.9	1.2	0
Ethanol-amine	14	3.7	2.2	9	0	4	0.7	388	10	1.1	5.7	4
Acetic acid	11	1.5	6.6	7	0	3	0.9	149	8	0.8	8.1	2
Propionic acid	8	2.6	1.4	9	0	3	0.5	242	5	1.6	2.9	0
Valeric acid	22	4.1	6.7	10	0	5	0.7	21	17	1.3	13	13
Lactose	149	5.7	25	11	0	13	0.2	478	147	0.5	431	144

Note. SMA = Steiglitz-McBride algorithm.

decomposition stages is provided in Figure 1. Decimation filters considered here are Coiflets (Coif5). Decomposition is obtained as follows[23]:

$$r^{q,s}(m) = \sum_{n=1}^{N'} \zeta'_n p'^m_n + v^{q,s}(m) \tag{11}$$

for $m = 0, 1, \cdots, \hat{M}-1$, where M' is the number of subband samples. $N' \leq N$ is the number of cisoids in the subband (q,s). ζ'_n and p'^m_n are subband counterparts of the fullband parameters ζ_n and p^m_n in the (q,s)th subband. Once the subband estimates $\hat{p}'^m_n$ are obtained, they can be converted to fullband using

$$\hat{p}^m_n = (\hat{p}'^m_n)^{1/2q} e^{j\pi\frac{2s+1}{2^q+1}} \quad \text{for} \quad \omega\in[0,\pi], \tag{12}$$

$$\hat{p}^m_n = (\hat{p}'^m_n)^{1/2q} e^{j\pi(\frac{2s+1}{2^q+1}-1)} \quad \text{for} \quad \omega\in[-\pi,0]. \tag{13}$$

Fullband estimates of $\hat{\zeta}'_n$ are computed using 9. Since the filters are not ideal, they will overlap as depicted in

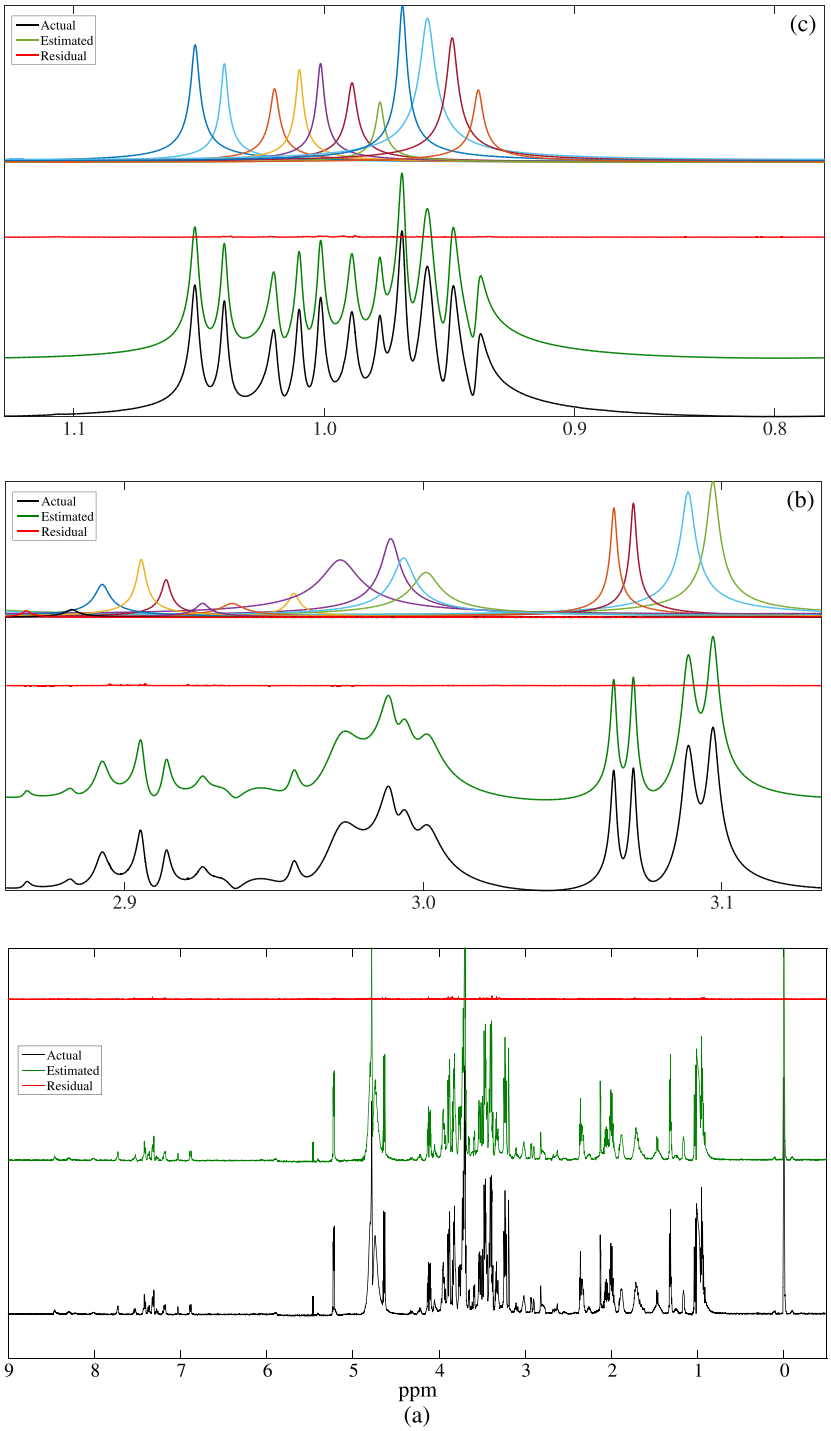

FIGURE 5 (a) Quantitation results achieved by the proposed algorithm for a representative fermentation broth in 90% H_2O. Proposed algorithm took 12.4 s compared with 3.5 min taken by complete reduction to amplitude frequency table[11] and does not require regions of interest to be specified. (b, c) Expansion of quantitation results displaying overlaid component spectra (from bottom to top, experimental, estimated, residual, and component spectra)

Figure 1. As a result, frequency components may appear in several bands. In order to uniquely determine to which subband a component belongs, the following criteria can be applied to select the correct band[17]:

$$\max|S_{q,s}(e^{j\hat{\omega}_i^{'}})|, \tag{14}$$

where $\hat{\omega}_i^{'}$ is the frequency chosen in the (q,s)th such that it had the highest amplitude in that band's spectrum $S_m(e^{j\hat{\omega}_i^{'}})$ as compared with the other bands in which it appeared. Finally, to make the subband decomposition adaptive, a BIC-based stop criteria using the frequency domain residual of the fullband signal can be adopted.[18]

6 | ALGORITHM

The entire process is summarized as an algorithm as follows:

1. Check each subband for poles according the BIC criteria in 10.
2. If no poles present, halt.
3. If poles present, extract parameters using SMA and decompose further.
4. Repeat steps from (1) to (3) until a global halt is reached.
5. Map estimated $p_n^{'m}$ to fullband values using 12 and 13.
6. Employ 9 to estimate $\hat{\zeta}_n$.

7 | EXPERIMENTAL RESULTS

In this section, we present experimental results obtained by the application of the proposed algorithm to real-time FID data. The data were acquired from a Magritek Spinsolve®SPA218 1.4.1.2717 system. The chemical sample used was ethanol. Its ^{1}H NMR spectrum—including multiplicity of spectral peaks due to the spin-spin coupling—can be completely characterized by its chemical structure.[24] The FID signal was used "as-is" without any preprocessing, for example, filtering, baseline correction or noise removal. The filters employed for subband processing were the fifth-order "Coiflet" filters due to their sharp cutoff, flat magnitude response, and relatively small group delay.[25]

Figure 2 shows the recorded FID signal for ethanol. Figure 3 depicts a comparison of spectrum of this signal with those estimated by the proposed algorithm and the SMA. Clearly, the latter is unable to capture all peaks of the ethanol spectrum. By contrast, the proposed algorithm successfully captures all of them; it was able to do so at the decomposition depth of 9 that took 2.8 s for a signal length of 32,768 samples on an Intel 3.40 GHz CPU. The execution time of the proposed algorithm is comparable with 1.9 s taken by the SMA, which missed 10 out of 14 spectral peaks.

Also in Figure 3, though no visible difference is observed between the true and estimated spectra at the given scale, Cramer-Rao Lower Bound (CRLB) analysis predicts some variance between the true and estimated peak amplitudes. This can be explained by considering the analytical CRLB derived for the damped cisoid

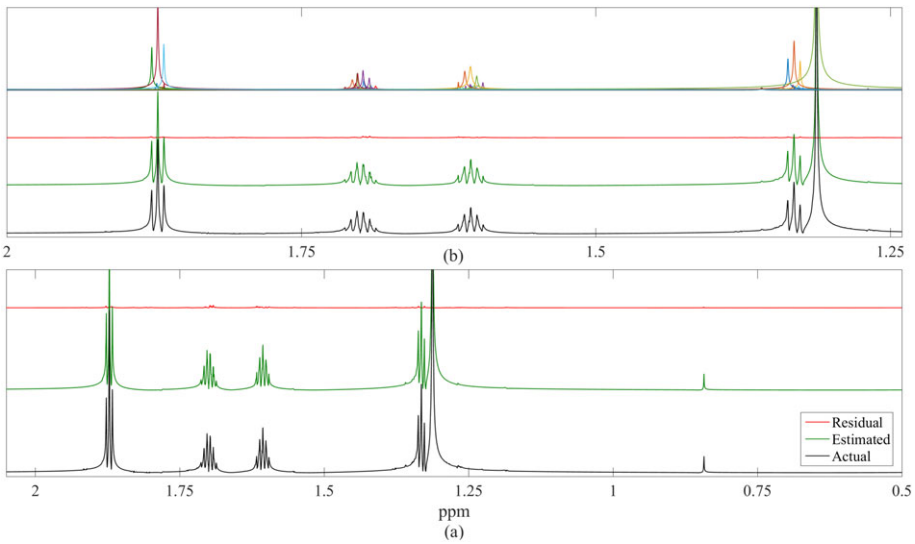

FIGURE 6 (a) Quantitation results obtained by the proposed algorithm for valeric acid (experimental, estimated, and residual spectra). (b) Expansion of quantitation results displaying overlaid component spectra (from bottom to top, experimental, estimated, residual, and component spectra)

model.[26,27] The results show that—out of all model parameters—complex amplitudes have the highest associated error variance. As a result, a difference in the amplitudes of true and estimated peaks would be observed. This variance cannot be reduced below the limit set by the CRLB.

For further elucidation, we also compared the spectra of numerous other chemical compounds with those reconstructed by the proposed algorithm, the SMA and the LCModel—a software for automatic quantitation of ^{1}H NMR data and freely available at http://s-provench-er.com. Each time, a different FID signal was selected and processed, and its spectrum was generated for comparison. Results are presented for the execution time, depth of decomposition for the proposed algorithm, mean squared error (MSE) between the reconstructed spectra, the number of peaks estimated in excess and the number of peaks missed, if any. The results are depicted in Table 1. These results again confirm our earlier observations; first that the execution time incurred by the proposed algorithm is comparable with the SMA/LCModel and second that the number of peaks missed by the SMA/LCModel becomes quite significant as the FID signal grows in complexity.

This is particularly evident in the case of the FID signal for *lactose*, a disaccharide sugar found in milk, and an extremely challenging signal for a quantitation method due to a very large number of signal samples (>30,000), a large number of resonances (>100), a very large dynamic range (90 dB), and an extremely challenging noise floor (with many resonances lying well below the noise). The ^{1}H spectrum for lactose is depicted in Figure 4.

Out of 149 peaks, the SMA and the LCModel were able to capture only two and six peaks, respectively. On the other hand, the proposed algorithm was able to resolve all 149 spectral peaks. This included the smallest resonance peak that, as also depicted in Figure 4, was almost 35,000 times smaller than the largest peak in the spectrum. The proposed algorithm was able to do so in just 5.7 s. The MSE achieved was 2.5×10^{-5}.

Finally, performance of the proposed algorithm is compared with CRAFT, a Bayesian algorithm for quantitative NMR mixture analysis.[11] The CRAFT analysis for 28 regions of interest of ^{1}H NMR spectrum of a representative fermentation broth in 90% H_2O took approximately 3.5 min. The proposed algorithm, on the other hand, took only 12.4 s to process the entire spectrum and, contrary to CRAFT, did not require the regions of interest to be specified. A close agreement between the true and estimated spectra is observed in Figure 5. A similar close agreement is demonstrated in Figure 6 for valeric acid.

8 | CONCLUSION

In this article, a subband SMA has been proposed for fast, accurate, and automatic estimation of the parameters of an FID signal for chemical spectroscopy. The proposed algorithm is more accurate in estimating the FID signal parameters when compared with the alternative parametric methods and has a similar speed.

ORCID

M.A.R. Anjum http://orcid.org/0000-0002-6733-3876

REFERENCES

[1] S. Umesh, D. W. Tufts. *IEEE Trans. Signal Process.* **1996**, *44*(9), 2245.

[2] G. L. Bretthorst. *Concepts Magn. Reson. Part A* **2005**, *27*(2), 73.

[3] A. A Istratov, O. F Vyvenko. *Rev. Sci. Instrum.* **1999**, *70*(2), 1233.

[4] E. Laue, J. Skilling, J. Staunton. *J. Magn. Reson. (1969)* **1985**, *63*(2), 418.

[5] H. Gesmar, J. J Led. *J. Magn. Reson. (1969)* **1988**, *76*(1), 183.

[6] S. Vanhuffel, H. Chen, C. Decanniere, P. Vanhecke. *J. Magn. Reson., Ser. A* **1994**, *110*(2), 228.

[7] P. Stoica, T. Sundin. *J. Magn. Reson.* **2001**, *152*(1), 57.

[8] M. Björk, D. Zachariah, J. Kullberg, P. Stoica. *Magn. Reson. Med.* **2016**, *75*(1), 390.

[9] V. A Mandelshtam. *Prog. Nucl. Magn. Reson. Spectrosc.* **2001**, *38*(2), 159.

[10] H. Barkhuijsen, R. De Beer, W. Bovee, D. Van Ormondt. *J. Magn. Reson. (1969)* **1985**, *61*(3), 465.

[11] K. Krishnamurthy. *Magn. Reson. Chem.* **2013**, *51*(12), 821.

[12] Y. Hua, T. K. Sarkar. *IEEE Trans. Acoust. Speech Signal Process.* **1990**, *38*(5), 814.

[13] S. Ye, E. Aboutanios. Efficient peak extraction of proton NMR spectroscopy using lineshape adaptation, in 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, Florence, Italy **2014**, 5661.

[14] E. Aboutanios, D. S. Thomas, J. M. Hook, C. Cobas. *J. Magn. Reson.* **2017**, *282*, 62.

[15] P. Stoica, P. Babu. *Digital Signal Process.* **2013**, *23*(5), 1565.

[16] J.-B. Poullet, D. M. Sima, S. Van Huffel. *J. Magn. Reson.* **2008**, *195*(2), 134.

[17] A. Tkacenko, P. Vaidyanathan, Sinusoidal frequency estimation using filter banks, in Acoustics, Speech, and Signal Processing, 2001. Proceedings (ICASSP'01). 2001 IEEE International Conference on, 5, IEEE, Salt Lake City, UT, USA, USA **2001**, 3089.

[18] E. H. Djermoune, M. Tomczak, An adapted filter bank for frequency estimation, in Signal Processing Conference, 2004 12th European, IEEE, Vienna, Austria **2004**, 2171.

[19] P. Stoica, Y. Selen. *IEEE Signal Process Mag.* **2004**, *21*(4), 36.

[20] P. Stoica, R. L. Moses, *Spectral Analysis of Signals*, 452, Pearson Prentice Hall Upper Saddle River, NJ **2005**.

[21] M. Tomczak, E. H. Djermoune. *J. Magn. Reson.* **2002**, *158*(1), 86.

[22] P. Stoica, T. Soderstrom. *IEEE Trans. Autom. Control* **1981**, *26*(3), 712.

[23] E. H. Djermoune, M. Tomczak, P. Mutzenhardt. *J. Magn. Reson.* **2004**, *169*(1), 73.

[24] R. S. Reid. *J. Chem. Educ.* **1989**, *66*(4), 344.

[25] J. S. Walker, *A Primer on Wavelets and their Scientific Applications*, CRC press, New York Washington DC **2008**.

[26] T. Wigren, A. Nehorai. *IEEE Trans. Signal Process.* **1991**, *39*(4), 1017.

[27] S. Cavassila, S. Deval, C. Huegen, D. Van Ormondt, D. Graveron-Demilly. *J. Magn. Reson.* **2000**, *143*(2), 311.

How to cite this article: Anjum MAR, Dmochowski PA, Teal PD. A subband Steiglitz-McBride algorithm for automatic analysis of FID data. *Magn Reson Chem*. 2018;56:740–747. https://doi.org/10.1002/mrc.4723