Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Magnetic Resonance 163 (2003) 192–197
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Communication

PowerSlicing

Søren Balling Engelsen* and Rasmus Bro

Centre for Advanced Food Studies, Department of Dairy and Food Science, Food Technology, The Royal Veterinary and Agricultural University, Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark

Received 13 December 2002; revised 24 March 2003

Abstract

Recently, a new technique for unique non-iterative multi-exponential fitting of time domain NMR data was proposed. The method was termed SLICING, because an intrinsic part of the method consisted of taking different parts (slices) of the original matrix data and rearranging the slices into a three-way box of data. Subsequently, a directly calculated model of this box provided T_2-estimates and corresponding amplitudes. The most critical part of this method is the choice of how to slice the original data. In this paper, a new general scheme for this slicing is proposed which (1) is shown to provide more accurate T_2-estimates and (2) leads to a significant speed improvement compared to earlier approaches. The method is called POWERSLICING, because it takes slices of lag 2^x ($x = 0, 1, \ldots, N$) where $2^N \leqslant J/2$ and J is the number of bins on the time axis. This approach ensures a reasonably high amount of direct constraints and an appropriate representation of both short and long time decays in the decomposition.

Keywords: Time domain NMR; Multi-exponential curve resolution; DECRA; Slicing; Transverse relaxation

1. Introduction

Time domain NMR (TD-NMR) data are often analyzed using multi-exponential decomposition, which applies non-linear and iterative exponential curve-fitting algorithms to find the underlying pure exponentials. Provided simultaneous analysis of several NMR transverse relaxation curves we have recently described an alternative non-iterative and rapid technique for curve resolution namely the SLICING technique [1]. The SLICING technique pseudo-upgrades the data to become tri-linear (residual magnetization is described as a function of two variables) which in turn facilitates some unique advantages offered by application of so-called tri-linear mathematical models. The method is based on the fact that two different time "slices" of a given multi-exponential decay curve consist of the same underlying qualitative features (characteristic decay times), but in a new linearly related combination of quantities (concentrations or magnitudes), utilizing the linear relationship between exponentials

$$\exp\left(\frac{-t}{T_{2n}}\right) \propto \exp\left(\frac{-t+\Delta t}{T_{2n}}\right). \tag{1}$$

In the simplest case a relaxation curve can be translated one data point, called *lag* 1, and added in a new direction called *slab* (*slab* 2), creating a data array (three-dimensional matrix) with the dimension two in the *slab* direction and the dimension $N-1$ in the *lag* direction (see Fig. 1). This (most basic) type of slicing corresponds to the original proposed DECRA slicing (direct exponential curve resolution algorithm [2]). Mathematically, the slicing operation can be described as follows. Let the matrix $\mathbf{X}$ ($I \times J$) contain the elements x_{ij} where i refers to a sample (row) and j to a time (column). Then measured data are simply expressed as x_{ij}, $i = 1, \ldots, I$; $j = 1, \ldots, J$. The above three-dimensional array can be described by its elements y_{ikm}, $i = 1, \ldots, I$; $k = 1, \ldots, J-1$; $m = 1, 2$. where $y_{ik1} = x_{ik}$ and $y_{ik2} = x_{i(k+1)}$.

The idea of "cutting" data into a number of overlapping slices has given rise to the name selected for this approach: SLICING. If this operation is performed on a series of multi-exponential decay curves, it is possible to obtain a tri-linear structure that can be analyzed by, for example, PARAllel FACtor analysis, PARAFAC [3]. The result of this procedure will ideally be exactly the

* Corresponding author. Fax: +45-3528-32-45.
E-mail address: se@kvl.dk (S.B. Engelsen).

1090-7807/03/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S1090-7807(03)00125-3

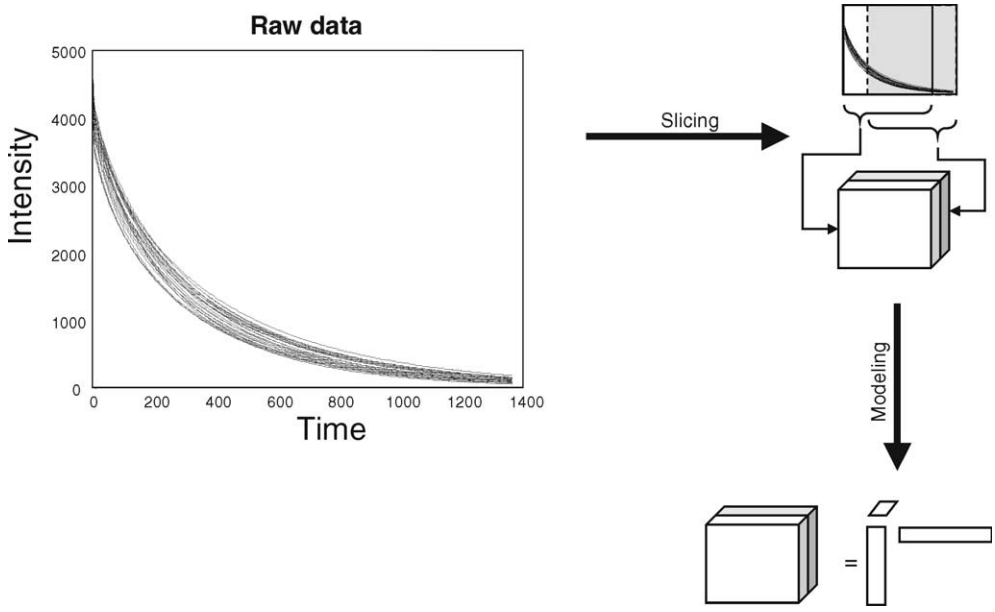


Fig. 1. The principle of slicing. A set of NMR relaxometry profiles (left figure) is held in a matrix with each row corresponding to one measurement. This matrix is split so that the leftmost columns are in one table and the rightmost in another (upper right figure). These slices are generally overlapping, and when the *lag* is one, the overlap is almost complete. The left slice has the original columns 1 to $J-1$ and the right slice has columns 2 to J. The tables are put behind each other resulting in a three-way array (residual magnetization as a function of two time variables) which can be modeled using PARAFAC (lower right figure).

same as that of a discrete exponential-fitting algorithm, fitting common characteristic time constants to a series of samples [1]. However, in contrast to the curve fitting approach that is based on a hard modeling to a functional form, the SLICING model is built on the use of highly redundant data where it is the context (amongst the slices) that determines the precise meaning. Extensive simulations of TD-NMR data revealed that the SLICING approach was comparable, but not superior to a robust classical numerical approach. However, the SLICING approach has proven to be practically important, because this algorithm utilizing highly redundant information and requiring no initial value guesses provides rapid, non-iterative and unique solutions with perfect mono-exponential loadings. The dramatically improved speed (independent of number of components extracted) and improved diagnostics (visual or numerical validation of the appropriateness of the solution) of the SLICING algorithm are its best attributes.

One potential problem with the SLICING algorithm is the multiple choice situation of the two meta-parameters *lag* and *slab* which in some cases significantly influences the calculated T_2-estimates. In our first report on the SLICING algorithm, we investigated a number of different choices including DECRA [2] slicing, a non-redundant slicing scheme and a meta-parameter optimization approach. As exemplified in Table 1, the accuracy of the time-constant estimates from a DECRA model and an optimized SLICING model may differ considerably. Evidently the optimized SLICING model provides the most accurate results, but unfortunately the optimized SLICING approach eliminates the speed advantage of the SLICING methods. For this reason a SLICING scheme that a priori is known to provide near-optimal results is greatly desired. During theoretical investigations of our SLICING algorithm we developed a slicing scheme which performs near optimal for TD-NMR relaxation curves and similar multi-exponential decay curves and which is similar in accuracy to the optimized SLICING model and which is superior in algorithmic speed, the latter being a good indication of an optimal model.

Table 1
T_2-estimates for two randomly selected theoretical sets (30 samples, 0.5% noise level) of relaxation curves using DECRA and optimized SLICING, respectively, as well as for the proposed method: POWERSLICING

T_2 (ms)	20	40	80	160
DECRA	21	46	107	317
	14	30	107	695
Optimized SLICING	20	40	82	163
	20	44	94	169
POWERSLICING	20	40	80	159
	20	41	82	160

The upper row shows the true T_2 values.

2. PowerSlicing

The purpose of data slicing is to indirectly impose exponential constraints by overlapping different parts of

the signal. The goal is to use a data slicing method that is optimal with respect to the accuracy of the solutions provided. In practice, the true parameters are not known and therefore the slicing method cannot be optimized specifically for each dataset, unless some derived statistical model quality measure is used (as is the case for Optimized SLICING). Even with a derived statistical quality measure, the globally optimal solution to the data slicing meta-parameter problem cannot be known without exploiting the whole meta-parameter space systematically. This is not realistic and will eliminate any advantages that the SLICING method may have.

Several aspects of a specific slicing scheme will affect the model quality.

1. An attractive representation with respect to SLICING is a representation where the underlying decays are as well spread as possible in both variable directions. Mathematically, the selectivity of each exponential should be as high as possible, which is obtained by using short *lags* for fast decays and long *lags* for slow decays.
2. The more (valid) constraints (*lags* and *slabs*) imposed, the smaller the variance of the estimated parameters will be.

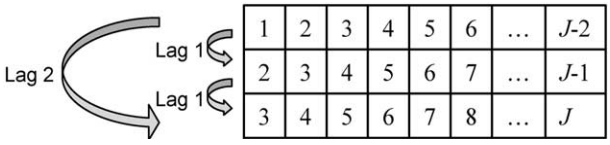
The new, generally applicable SLICING model that we propose is based on the fact that a slicing scheme can be developed with non-equidistant lagging, as was also investigated in the previous investigation [1]. It is based on the idea of obtaining a simple lagging scheme that imposes many direct exponential constraints amongst the variables and the fact that the second 'relaxation' direction provides maximum contrast between possible individual relaxation components while keeping the number of slabs low. The new SLICING model uses a lagging function of the form 2^x $(x = 0, 1, \ldots, N)$ for which reason it is named POWERSLICING. Using this functional form, the resulting data matrix from one random TD-NMR profile is shown in Fig. 2. The advantage of this lagging scheme is outlined below.

In order to demonstrate how POWERSLICING works, it is instructive to examine a small example. Let one decay curve be represented by a vector **x** with index elements numbered $1, 2, 3, \ldots, J$. When this curve is sliced once with a *lag* of one (DECRA), the resulting two-slab matrix can be written as

1	2	3	4	5	6	$\cdots$	$J-1$
2	3	4	5	6	7	$\cdots$	J

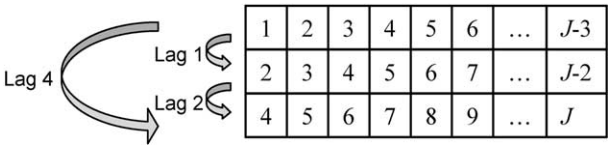
For more than one sample, a three-way array is obtained. When modeling these data with SLICING, the model imposes the constraint that both rows can be described by the same exponential profiles (T_2's) up to a scalar difference. Thus, for a one-component model, SLICING defines that row one shown above can be approximated by a $(J-1)$-vector **b** and therefore row two as $d\mathbf{b}$ where d is a constant. Thus, by *lagging* one, SLICING *directly* imposes the exponential feature explained in Eq. (1) only for *adjacent* points in the TD-NMR profile.

If a new *slab* is added which starts at element 3, we obtain the following three-*slab* model:



Compared to the two-*slab* representation, the addition of one more slice introduces two explicit constraints. Row 1 versus row 3 imposes the exponential constraint for points with *lag* 2 (e.g., element 1 versus 3 and element 2 versus 4). The second constraint imposed between row 2 and row 3, however, is not *active*, because it is identical to the constraint between row 1 and row 2. As for the difference between row 1 and row 2, row 2 and row 3 only implies the exponential constraint between adjacent points. Therefore, the constraint leads to overly redundant (thus computationally costly) data representation without adding new restrictions.

Consider instead an alternative lagging (2^x or POWERSLICING) where the third row starts at element 4 (2^2):



For this representation, the two new constraints imposed are between points with *lag* 2 (row 2 versus row 3) and lag 3 (row 1 versus 3). As can be seen, this lagging introduces truly new constraints. Adding such new constraints is helpful in creating diversity in the row-mode, which is one of the premises for successful SLICING modeling.

Continuing the POWERSLICING approach, a new row starting at element 8 (2^3) again introducing only truly new direct constraints:

1	2	3	4	5	6	$\cdots$	$J-7$
2	3	4	5	6	7	$\cdots$	$J-6$
4	5	6	7	8	9	$\cdots$	$J-4$
8	9	10	11	12	13		J

From the example above it follows that an effective and generally applicable lagging scheme can be obtained by starting at elements 2^0, 2^1, 2^2, etc. This ensures that the exponential aspect is imposed as much as possible in POWERSLICING, using as few lagged slabs as possible.

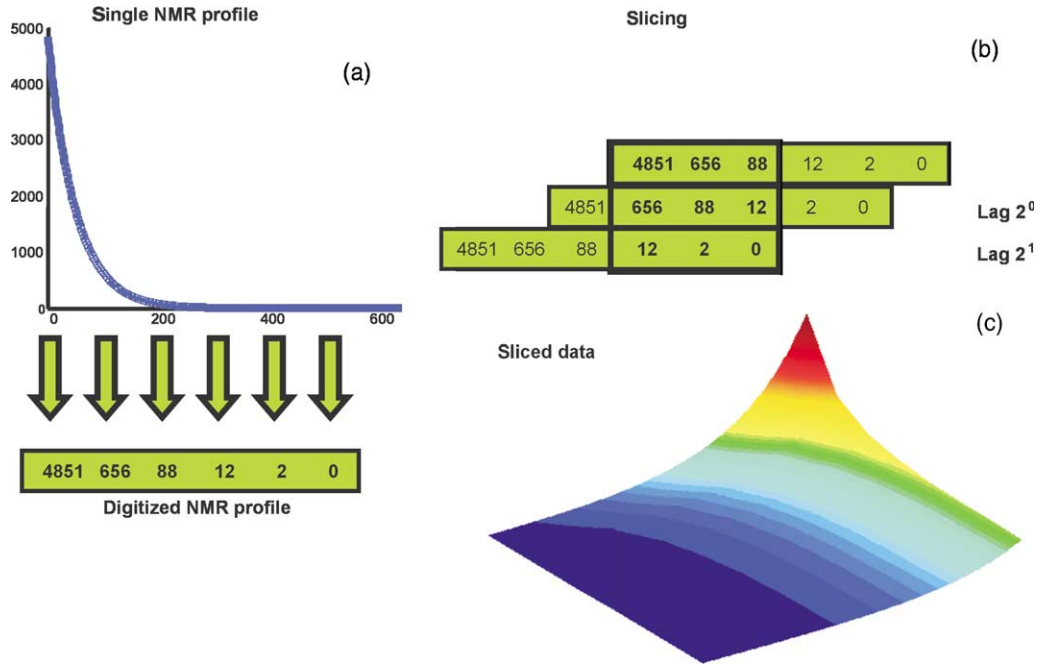


Fig. 2. Illustration of the digitized relaxometric data. (a) Digitizing the relaxometric profile, (b) the lagging scheme of the three first slices in PowerSlicing, and (c) an example of the resulting decay landscape of a single multi-exponential NMR profile.

PowerSlicing does not take into account all possible constraints. For example, the first exponential constraint that is not directly imposed is between variable J and $J-5$. However, this constraint is indirectly taken into account via the constraints between variables J and $J-4$ and between $J-4$ and $J-5$.

Having defined the PowerSlicing lagging procedure, the second problem is then the choice of the number of slabs. It is possible to have lags as high as $J/2$, whereas higher lags would be beyond the end of the curve. Will there be an advantage in taking fewer slabs than the maximal possible number? It is conjectured that there is not. Fewer slabs may introduce less correlation structure in the residuals and will further increase the variance of the estimates, because more structure is imposed. It is therefore suggested that the number of slabs taken always be as high as possible. In Section 3, this choice is verified empirically.

3. Results

To test the PowerSlicing technique we applied it to a simulated NMR dataset of 30 samples with 2048 data points containing four underlying exponential components interspaced with a factor of two and added 0.5% (of total sum of squares of the data) random noise. This was repeated 3000 times using different random noise in each case. First, the number of slabs to be used was determined. Fig. 3 shows the result of using 2 (DECRA), 3, 6, and 11 slabs of the 2^x lagging function,

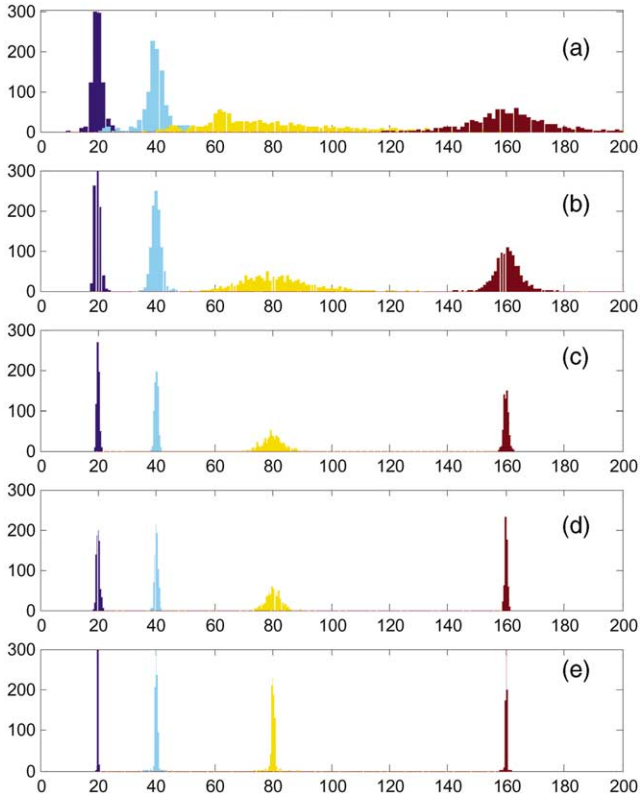


Fig. 3. Dispersion of T_2 values found by PowerSlicing (2^x) as a function of number of slabs ($n = 1, 2 \ldots, x$). (a) Slicing with 2 slices using the lags 0 and 2^0 (DECRA), (b) Slicing with 3 slices using the lags 0, 2^0, and 2^1, (c) Slicing with 6 slices using the lags 0, 2^0, 2^1, 2^2, 2^3, and 2^4, (d) Slicing with 11 slices using the lags 0, 2^0, 2^1, 2^2, 2^3, 2^4, 2^5, 2^6, 2^7, 2^8, and 2^9 (PowerSlicing), and (e) MatrixFit.

i.e., the last slab will be lagged by 1024 data points. Two important observations may be deduced from the figure. Most importantly, the dispersion of the T_2-estimates is monotonically reduced with the number of slabs used, the maximum *slab* (11) PowerSlicing being the best. This type of data experiment suggests that the optimal number of *slabs*, N, for the 2^x lagging function is $2^N \leqslant J/2$, where J is the number of bins on the time axis. This maximum model will from now on be labeled PowerSlicing. Second, PowerSlicing turned out to be the most efficient algorithm, which on average is a factor of two faster than the *slab* 2 (or DECRA) model and a factor of 6 faster than the traditional approach implemented in the MatrixFit algorithm [1], indicating a more optimal and convergent model.

Fig. 4 compares the performance of PowerSlicing to a traditional numerical approach MatrixFit [1], the previously introduced optimized Slicing [1] and DECRA [2]. From the figure it is observed that the dispersion of the calculated T_2-estimates by PowerSlicing is comparable and even somewhat more narrower than the results from the optimized Slicing; however, PowerSlicing remains a little less accurate in the T_2-estimates when compared to the MatrixFit approach. When examining the average algorithmic speed, the PowerSlicing outperforms the optimized Slicing by a factor of approximately 300 and the MatrixFit approach by a factor of 6.

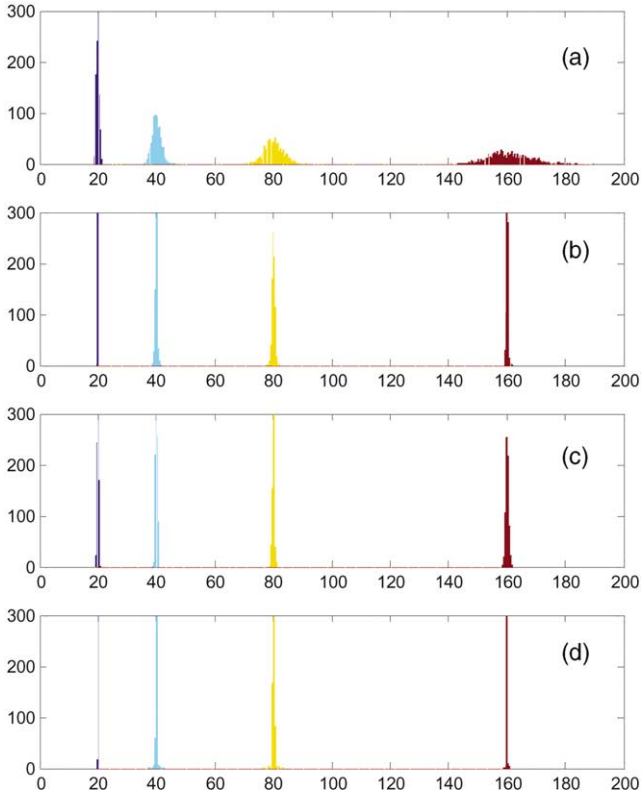


Fig. 4. Dispersion of T_2 values found by (a) DECRA, (b) SlicingOpt, (c) PowerSlicing, and (d) MatrixFit. Note that there is no significant bias in any of the estimates.

Table 2
Linear regression correlation and cross-validated prediction performance between fat content and the concentration vector best describing the fat content

Algorithm	T^2 (ms)	# NoComp	*r2*	RMSECV	CPU
Multi-exponential fitting	⟨247⟩	2	0.969	0.650	25.3
	⟨424⟩	3	0.857	1.391	61.8
MatrixFit	284	2	0.964	0.696	3.3
	420	3	0.975	0.588	7.4
DECRA	295	2	0.964	0.699	1.7
	182	3	0.969	0.651	2.2
PowerSlicing	306	2	0.965	0.691	1.1
	358	3	0.976	0.572	1.2

RMSECV: Root Mean Square Error of Cross-Validation.

To test the validity of the PowerSlicing approach we applied it to a data set that consists of TD-NMR measurements on 47 samples of minced meat with a total fat content ranging from 1.2 to 15% (w/w) [4]. In this experiment the standard pulsed field gradient stimulated echo experiment was followed by the CPMG 180° pulse train. The samples were measured at 55 °C to ensure liquid fat phase and a total of 2048 even echoes were acquired using a τ of 500 μs. Prior to analysis the data were phase corrected using principal phase correction (PPC) [1], a necessary prerequisite for sound results. The results of PowerSlicing on this data set compared to the three above-mentioned algorithms are presented in Table 2. From the table it is observed that the quantitative performance of the one-dimensional multi-exponential fitting deteriorates when using more than two components, while the performance of the two-dimensional data technology approaches is improved when using three components. Only insignificant differences are observed between MatrixFit and the PowerSlicing, whereas DECRA appears to perform inferiorly. Moreover, it is observed that the T_{2X} time constant for the fat component is relatively more stable in the PowerSlicing approach when increasing the model complexity from two to three components, albeit a considerable increase is observed. The CPU index clearly demonstrates the main benefit of the PowerSlicing.

4. Conclusion

We have previously proposed a new data technique for simultaneous multi-exponential fitting of TD-NMR relaxation curves called Slicing. The greatest virtues of the Slicing method are that it is non-iterative and fast. It provides unique solutions (for a predetermined number of components) in the mathematical sense and it

requires no initial guesses. It also provides exploratory tools that can help assess the validity of the solution. The disadvantage of the originally proposed SLICING method was the lack of a generally applicable data slicing scheme which made optimization of meta-parameters necessary, which in turn eliminated the speed advantage. In this study, we propose a generally applicable near-optimal data slicing method called POWERSLICING which has proven to provide accurate and rapidly calculated T_2-estimates for both theoretical and experimental data. The new data slicing method provides considerably more robust estimates than any previously proposed slicing scheme and we recommend that all ongoing data slicing experiments use the new technique.

Acknowledgments

The authors gratefully acknowledge support from the Centre for Advanced Food Studies, the Danish Agricultural and Veterinary Research Council (the project entitled: Centre for Critical Quality Attribute Determination in Muscle Foods), and the Danish Technical Research Council (Project 1179) for financial support.

References

[1] H.T. Pedersen, R. Bro, S.B. Engelsen, Towards rapid and unique curve resolution of low-field NMR relaxation data: trilinear SLICING versus two-dimensional curve fitting, Journal of Magnetic Resonance 157 (1) (2002) 141–155.
[2] W. Windig, B. Antalek, Direct exponential curve resolution algorithm (DECRA): a novel application of the generalized rank annihilation method for a single spectral mixture data set with exponentially decaying contribution profiles, Chemometrics and Intelligent Laboratory Systems 37 (1997) 241–254.
[3] R. Bro, PARAFAC. Tutorial and applications, Chemometrics and Intelligent Laboratory Systems 38 (1997) 149–171.
[4] H.T. Pedersen, H. Berg, F. Lundby, S.B. Engelsen, The multivariate advantage in fat determination in meat by Bench-top NMR, Innovative Food Science and Emerging Technologies 2 (2) (2001) 87–94.