Shifted factor analysis—Part III: *N*-way generalization and application

Sungjin Hong* and Richard A. Harshman

University of Western Ontario, London, Ontario N6A 5C2, Canada

The 'quasi-ALS' algorithm for shifted factor estimation is generalized to three-way and *n*-way models. We consider the case in which mode A is the only shifted sequential mode, mode B determines shifts, and modes above B simply reweight the factors. The algorithm is studied using error-free and fallible synthetic data. In addition, a four-way chromatographic data set previously analyzed by Bro *et al.* (*J. Chemometrics* 1999; 13: 295–309) is reanalyzed and (two or) three out of four factors are recovered. The reason for the incomplete success may be factor shape changes combined with the lack of distinct shift patterns for two of the factors. The shifted factor model is compared with Parafac2 from both theoretical and practical points of view. Copyright © 2003 John Wiley & Sons, Ltd.

1. INTRODUCTION

The three-way shifted factor problem arises in various contexts, for example in electrophysiology [1] and some chemometric problems [2,3]. We previously described the two- and three-way shifted factor models [4] and developed a 'quasi-ALS' algorithm to fit the two-way model [5]. In the first section of this paper we extend the 'quasi-ALS' approach to three- and higher-way data.

As in the two-way case, the three-way multilinear expression does not allow for position-shifted factors underlying the data. One way to resolve the three-way shifting problem, first proposed by Bro [2,6], is to use Parafac2. One can directly fit Parafac2 to the profile data [7] or indirectly fit Parafac2 to derived second-moment type data (e.g. cross-product, covariance or correlation matrices) [8]. Typically, three-way Parafac2** is written in terms of the cross-products computed from the *k*th (frontal) slab of an $I \times J \times K$ data array. However, to facilitate the shifted factor generalization, we will write it in terms of the *j*th lateral slab, as

$$\mathbf{X}_j'\mathbf{X}_j = \mathbf{C}\langle\mathbf{b}_j\rangle\mathbf{\Phi}\langle\mathbf{b}_j\rangle\mathbf{C}' + \mathbf{E}_j$$
$$\mathbf{\Phi} = \mathbf{A}_j'\mathbf{A}_j \tag{1}$$

where $\mathbf{X}_j'\mathbf{X}_j$ is a $K \times K$ cross-product matrix at level *j* (*K* being the number of levels of mode C) and $\mathbf{\Phi}$ is a cross-product

matrix for $\mathbf{A}_j$. Using array notation (AIN) [9], the Parafac2 model in representative slice form corresponding to (1) is

$$y_{jKK'} = b_{jR}b_{jR'}c_{KR}c_{K'R'}\phi_{RR'}$$
$$\phi_{RR'} = a_{IjR}a_{IjR'} \tag{2}$$

The condition on $\mathbf{\Phi}$ stated in (1) or (2) implies the invariance of the factor relations in mode A across all levels of mode B, but not the invariance of mode A factor loadings themselves [8]. Bro uses this fact to relax the parameter invariance requirement in the position-shifted mode, thereby avoiding the position shifting problem.

However, the $\mathbf{\Phi}$ invariance requirement is met only when the position shift does not change the angles between sequential factors across the *J* slices. As Bro *et al.* [2] pointed out, one such case occurs when all sequential factors are shifted by the same amount at each of *J* slices; such identical shifting does not alter the off-diagonal elements in $\mathbf{\Phi}$, provided that the sequential factor loadings are negligible at both ends. If not, the values in $\mathbf{\Phi}$ will in general vary across mode B levels under the position shifting of factors.

When the size of shift is not the same for all factors, however, Parafac2 may fail to handle the position-shifting problem, since factorwise independent shifting generally causes $\mathbf{\Phi}$ to vary across the *J* slices. In other words, the independent shifting can cause *different* manifestations, across *J* slices, of otherwise the *same* factor angles. One may consider the factorwise identical shifting to be the shifting of the (non-error part of the) *observed* sequential data, and the factorwise independent shifting as the inde-

*Correspondence to: S. Hong, Department of Psychology, University of Western Ontario, London, Ontario N6A 5C2, Canada.
E-mail: shong2@uwo.ca
**Note that (1) becomes indirect-fit Parafac1 if the sequential factors are orthogonal, i.e. if $\mathbf{\Phi} = \mathbf{I}$.

pendent shifting of *latent* sequential factors [4,10]. Unlike Parafac2, SFA directly adjusts for the independent position shifting at the factor level.

To derive the shifted factor model, we start from the trilinear Parafac/Candecomp model [11,12]. By writing this model in 'representative slice' form [13] for lateral ($I \times K$) slices of the array, we get

$$\mathbf{X}_j = \mathbf{A}\langle \mathbf{b}_j \rangle \mathbf{C}' + \mathbf{E}_j \qquad (3)$$

where $\mathbf{A}$ and $\mathbf{C}$ are factor loading matrices for modes A and C respectively and $\mathbf{b}_j$ is the jth row in the mode B factor loading matrix $\mathbf{B}$; as usual, $\mathbf{E}_j$ represents residuals, in this case the jth slice of the $I \times J \times K$ residual array. As noted earlier, the use of lateral slices is unconventional but will be required to remain consistent with shifted factor mode conventions [4]. We can avoid deciding how to slice the data array if we use array notation [9], which gives a symmetric representation of the standard Parafac1 model of the full three-way array* as

$$x_{IJK} = a_{IR}b_{JR}c_{KR} + e_{IJK} \qquad (4)$$

The simplest shifted factor generalization of the trilinear model allows the factor positions to differ across the levels of only one mode. Let mode A be a sequential mode and let the sequential factors be independently shifted across levels of the shifting mode (mode B). Then SFA models slice j in a three-way data array $\underline{\mathbf{X}}$ as

$$\mathbf{X}_j = \mathcal{S}_j(\mathbf{A})\langle \mathbf{b}_j \rangle \mathbf{C}' + \mathbf{E}_j \qquad (5)$$

where the shifted version of $\mathbf{A}$ for slice j, $\mathcal{S}_j(\mathbf{A})$, is the same as in the two-way SFA model [5]. It is implied in (5) that the sequential factor profiles are fixed across all J slices before they are shifted at level j by the amounts specified in row j of the shift size matrix $\mathbf{S}$. That is, SFA actually estimates sequential factors that are invariant in shape across all levels of the other modes, but their position changes across the J mode B levels. Thus the SFA model (5) can be considered more restrictive than Parafac2 since it estimates the actual shape of the sequential factors and not just their relations. On the other hand, it is less restrictive than direct fit Parafac1 since it allows for changes in factor position.

Likewise the SFA model can be extended to n-way shifted data as

$$\mathbf{X}_{jl\cdots} = \mathcal{S}_j(\mathbf{A})\langle \mathbf{b}_j \rangle \langle \mathbf{d}_l \rangle \cdots \mathbf{C}' + \mathbf{E}_{jl\cdots} \qquad (6)$$

where $\mathbf{X}_{jl\cdots}$ is the $I \times K$ matrix at level $j, l, \ldots$ in the $I \times J \times K \times L \times \cdots$ n-way data array $\underline{\mathbf{X}}$.

In AIN, (5) can be written for the full three-way array as

$$x_{IJK} = a_{[I+s_{jr}]R}b_{JR}c_{KR} + e_{IJK} \qquad (7)$$

and (6) can be written for the full n-way array as

$$x_{IJKL\cdots} = a_{[I+s_{jr}]R}b_{JR}c_{KR}d_{LR}\cdots + e_{IJKL}\cdots \qquad (8)$$

The estimation of parameters for SFA models is difficult, because the shifting disturbs the multilinearity of factor variation. This interferes with attempts to estimate factor weights (e.g. $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ for model (5)) by regression

*The 'representative slice' version corresponding to (3) (i.e. sliced by levels of mode B) is simply $x_{IJK} = a_{IR}b_{jR}c_{KR} + e_{IJK}$.

methods, as would be done in an ALS (alternating least squares) approach. Here we describe some algorithms that deal with this problem. As in Part II [5], we simplify the exposition by only using the matrix approach (refer to Reference [14] for the AIN equivalents of all equations in the following sections).

2. THREE-WAY AND *N*-WAY GENERALIZATIONS OF QUASI-ALS ESTIMATION

We now focus on modifications for the three-way and n-way generalizations of the SFA models (5) and (6) in which the pattern of factor shifting differs across levels of only one mode (i.e. mode B or shifting mode), as in the two-way SFA model. The basic logic of mode A and B estimation remains the same as in the two-way case. The estimation of the mode(s) other than modes A and B is new but is similar to the ALS estimation in standard Parafac1. As a result, the three-way SFA (5) straightforwardly generalizes to the n-way SFA (6).

2.1. B and S estimation
The two mode B parameter sets $\mathbf{B}$ and $\mathbf{S}$ must be estimated in an interlocking fashion, as in the two-way case, except that $\mathbf{b}_j$ is now estimated by

$$\hat{\mathbf{b}}_j = \left[\left(\mathbf{C} \odot \mathcal{S}_j(\mathbf{A}) \right)' \right]^+ \mathbf{x}_j^{(IK \times 1)} \qquad (9)$$

where $\odot$ represents the Khatri–Rao product (or columnwise Kronecker product) [15]. Here $\mathbf{x}_j^{(IK \times 1)}$ is an $IK \times 1$ vectorized version of $\mathbf{X}_j$, ordered to conform with the generalized inverse of the combined fixed parameters.

As before, (9) implies that $\mathbf{s}_j$ is provided. If we assume that shifts are integral values and that their maximum value is $s_{\max}$, it is again possible to find the conditional global optimum by means of the Exhaustive Integer Search (EIS) procedure [5]. Equation (9) can be extended for the n-way SFA (6) as

$$\hat{\mathbf{b}}_j = \left[\left(\cdots \mathbf{D} \odot \mathbf{C} \odot \mathcal{S}_j(\mathbf{A}) \right)' \right]^+ \mathbf{x}_j^{(IKL\cdots \times 1)} \qquad (10)$$

where $\mathbf{x}_j^{(IKL\cdots \times 1)}$ is an $IKL \cdots \times 1$ vectorized version of the $I \times K \times L \times \cdots$ jth sub-block of the n-way data array, ordered so that it can be multiplied with the combined fixed parameters of the modes other than mode B.

2.2. A estimation
In order to undo the shifting in the data so that the alignment of the sequential mode levels is consistent between the data and $\mathbf{A}$, the unshifting must now be done on each *slice j* in the data. It is also necessary to separate the data in each *slice* into two parts before unshifting: one part is the variance contributed by factor r and the other is the variance contributed by the other factors.

To estimate the factor r contribution to slice j, we use

$$\hat{\mathbf{X}}_{j(r)} = \mathbf{X}_j - \mathcal{S}_j(\mathbf{A}_{(\sim r)})\langle \mathbf{b}_{j(\sim r)} \rangle \mathbf{C}'_{(\sim r)} \qquad (11)$$

for the three-way case, and this can be generalized for the n-way case as

$$\hat{\mathbf{X}}_{jl\cdots(r)} = \mathbf{X}_{jl\cdots} - \mathcal{S}_j(\mathbf{A}_{(\sim r)})\langle \mathbf{b}_{j(\sim r)} \rangle \langle \mathbf{d}_{l(\sim r)} \rangle \cdots \mathbf{C}'_{(\sim r)} \qquad (12)$$

Now that the $\hat{\mathbf{X}}_{j(r)}$ (or $\hat{\mathbf{X}}_{jl\cdots(r)}$) variance is isolated from the $\hat{\mathbf{X}}_{j(\sim r)}$ (or $\hat{\mathbf{X}}_{jl\cdots(\sim r)}$) variance, one can realign (or unshift) these J factor r contributions 'slicewise' in the three-way array $\hat{\underline{\mathbf{X}}}_{(r)}$ (or sub-blockwise in the n-way isolated data array), using the factor r shift values contained in $\mathbf{s}_r$ (column r of $\mathbf{S}$). This can be written as

$$\underline{\widetilde{\mathbf{X}}}_{(r)} = \overset{\hat{\mathbf{X}}_{j(r)}}{\underset{\mathbf{s}_r}{\mathcal{S}}}{}^{-1}\left(\hat{\underline{\mathbf{X}}}_{(r)}\right) = \overset{\hat{\mathbf{X}}_{j(r)}}{\underset{-\mathbf{s}_r}{\mathcal{S}}}\left(\hat{\underline{\mathbf{X}}}_{(r)}\right) \tag{13}$$

where the shifting operator $\mathcal{S}(\cdot)^{-1}$ inverse-shifts each of J lateral slices $\hat{\mathbf{X}}_{j(r)}$ in the three-way array $\hat{\underline{\mathbf{X}}}_{(r)}$ by the jth value in $\mathbf{s}_r$ or, equivalently, shifts each of J lateral slices in $\hat{\underline{\mathbf{X}}}_{(r)}$ by the jth value in $-\mathbf{s}_r$. Refer to the Appendix in Part I [4] for details of the shifting operation. Using the shorthand convention, (13) would be simply written as $\mathcal{S}_r^{-1}\left(\hat{\underline{\mathbf{X}}}_{(r)}\right)$.

Once the factor isolation and subsequent unshifting are done, one can unfold the three-way (or n-way) array $\underline{\widetilde{\mathbf{X}}}_{(r)}$ into an $I \times JK$ (or $I \times JKL\cdots$) matrix for use in the regression estimation of $\mathbf{a}_r$. Now $\mathbf{a}_r$ can be estimated as

$$\hat{\mathbf{a}}_r = \widetilde{\mathbf{X}}_{(r)}^{(I\times JK)}\left(\mathbf{c}_r' \otimes \mathbf{b}_r'\right)^{+} \tag{14}$$

where $\widetilde{\mathbf{X}}_{(r)}^{(I\times JK)}$ is an $I \times JK$ isolated, inverse-shifted and then unfolded data matrix for factor r and $\otimes$ is the Kronecker product. Equation (14) can be extended for the n-way case as

$$\hat{\mathbf{a}}_r = \widetilde{\mathbf{X}}_{(r)}^{(I\times JKL\cdots)}\left(\cdots \mathbf{d}_r' \otimes \mathbf{c}_r' \otimes \mathbf{b}_r'\right)^{+} \tag{15}$$

where $\widetilde{\mathbf{X}}_{(r)}^{(I\times JKL\cdots)}$ is an $I \times JKL\cdots$ isolated, unshifted and unfolded data matrix for factor r.

2.3. WLS for the ends of factor profiles

To deal with the complication caused by the cells that become empty after unshifting, we use the weighted least squares (WLS) approach as used in the two-way case. The WLS procedure for the augmented $\mathbf{A}$ in three-way SFA can be written as

$$\hat{a}_{ir} = \tilde{\mathbf{x}}_{i(r)}'^{(1\times JK)}\left[\left(\mathbf{c}_r' \otimes \mathbf{b}_r'\right) * \mathbf{w}_{i(r)}'^{(1\times JK)}\right]^{+} \tag{16}$$

The vector $\tilde{\mathbf{x}}_{i(r)}'^{(1\times JK)}$ is row i in the $I \times JK$ isolated, unshifted and unfolded data matrix $\widetilde{\mathbf{X}}_{(r)}^{(I\times JK)}$; and $\mathbf{w}_{i(r)}'^{(1\times JK)}$ is a $1 \times JK$ weighting vector containing 1s (for valid cells in the 'estimated' data) and 0s (for empty cells and so filled with zeros), which is row i in the $I \times JK$ weighting matrix $\mathbf{W}_{(r)}$ for factor r. The extended n-way version of (16) can be written as

$$\hat{a}_{ir} = \tilde{\mathbf{x}}_{i(r)}'^{(1\times JKL\cdots)}\left[\left(\cdots \mathbf{d}_r' \otimes \mathbf{c}_r' \otimes \mathbf{b}_r'\right) * \mathbf{w}_{i(r)}'^{(1\times JKL\cdots)}\right]^{+} \tag{17}$$

where $\tilde{\mathbf{x}}_{i(r)}'^{(1\times JKL\cdots)}$ and $\mathbf{w}_{i(r)}'^{(1\times JKL\cdots)}$ are row i respectively in the $I \times JKL\cdots$ isolated, unshifted and unfolded data matrix $\widetilde{\mathbf{X}}_{(r)}^{(I\times JKL\cdots)}$ and in the $I \times JKL\cdots$ weighting matrix $\mathbf{W}_{(r)}$ for factor r. Recall that the three-way (or n-way) data must first be augmented, as before, by adding the beyond-window levels, and hence the I in (16) and (17) is (the original I) + $2s_{\max}$.

2.4. C estimation

The estimation of $\mathbf{C}$ bears the closest resemblance to the ALS estimation in Parafac1. The only difference is that the columns of $\mathbf{A}$ must be properly shifted before multiplying them by $\mathbf{b}_j$ to obtain the fixed parameters. Unlike the estimation of

$\mathbf{A}$ and $\mathbf{B}$, one estimates all elements in $\mathbf{C}$ simultaneously, as in the ALS estimation of Parafac1. A properly shifted and combined set of fixed parameters appropriate for this simultaneous estimation can be obtained as

$$\widetilde{\mathbf{AB}} = \begin{bmatrix} \mathcal{S}_1(\mathbf{A})\langle\mathbf{b}_1\rangle \\ \mathcal{S}_2(\mathbf{A})\langle\mathbf{b}_2\rangle \\ \vdots \\ \mathcal{S}_J(\mathbf{A})\langle\mathbf{b}_J\rangle \end{bmatrix} \tag{18}$$

where $\widetilde{\mathbf{AB}}$ is an $IJ \times R$ partitioned matrix in which the columns of partition j are shifted by the values in $\mathbf{s}_j$ and then weighted by $\mathbf{b}_j$. Equation (18) can be extended for the n-way case as

$$\widetilde{\mathbf{ABD}}\cdots = \cdots\mathbf{D} \odot \begin{bmatrix} \mathcal{S}_1(\mathbf{A})\langle\mathbf{b}_1\rangle \\ \mathcal{S}_2(\mathbf{A})\langle\mathbf{b}_2\rangle \\ \vdots \\ \mathcal{S}_J(\mathbf{A})\langle\mathbf{b}_J\rangle \end{bmatrix} \tag{19}$$

where $\widetilde{\mathbf{ABD}}\cdots$ is an $IJL\cdots \times R$ partitioned matrix in which the columns of the innermost partition involving levels $\ldots, l$ and j are shifted by the values in $\mathbf{s}_j$ and then weighted by $\mathbf{b}_j$, $\mathbf{d}_l$ and so on. Given $\widetilde{\mathbf{AB}}$ (or $\widetilde{\mathbf{ABD}}\cdots$) and the $K \times IJ$ unfolded data matrix $\mathbf{X}^{(K\times IJ)}$ (or $\mathbf{X}^{(K\times IJL\cdots)}$), $\mathbf{C}$ can be estimated for the three-way case as

$$\hat{\mathbf{C}} = \mathbf{X}^{(K\times IJ)}\left[\left(\widetilde{\mathbf{AB}}\right)'\right]^{+} \tag{20}$$

and for the n-way case as

$$\hat{\mathbf{C}} = \mathbf{X}^{(K\times IJL\cdots)}\left[\left(\widetilde{\mathbf{ABD}}\cdots\right)'\right]^{+} \tag{21}$$

2.5. Fractional shift estimation

Once all parameters have converged in the EIS (Exhaustive Integer Search) stage, the quasi-ALS procedure shifts to the FLS (Fractional Line Search) stage, as in the two-way case [5]. The FLS procedure for the 'higher-way' cases is the same as for the two-way case. This is because we limit our current quasi-ALS algorithms to the cases where sequential factors shift differently only on one mode (mode B) and the other modes (above mode B) simply reweight the factors.

2.6. Modified procedures to accelerate estimation

As the number of data modes increases, the application of the current quasi-ALS procedure can become impractically slow. Therefore it is necessary to apply with caution a couple of the acceleration techniques described in the Appendix of Part II [5].

3. DATA ANALYSIS

We first tested the quasi-ALS procedure by applying the three-way SFA model (5) to error-free and fallible synthetic data. We then fit the four-way SFA to a four-way chromatographic data set previously analyzed by Bro *et al.* [2], in which factors were shifted along the elution time mode. All analyses used two-stage EIS (i.e. EIS1 and EIS2) and the periodic shift search (i.e. applied each of the first 10 iterations and then every 10th thereafter) in order to make the computation time practical (see the Appendix of Part II [5] for a
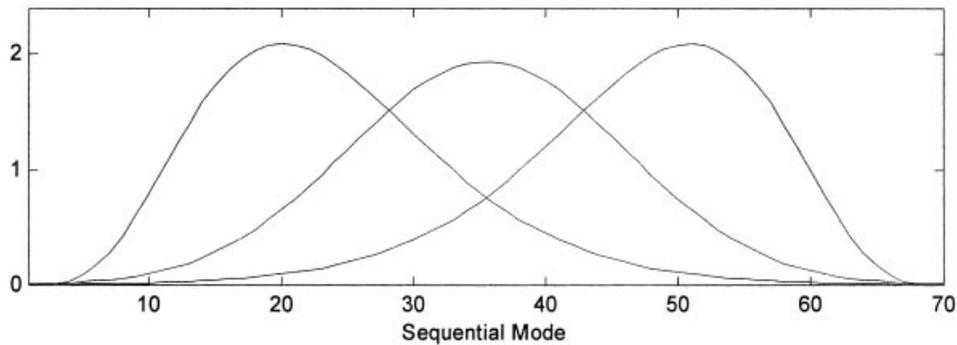
Figure 1. Sequential factors (type I). The middle curve is the normal density function of 70 evenly spaced standard deviates from -3.3 to 3.3. The positively skewed curve (left) is the chi-square density function of 70 evenly spaced chi-square values from 0.5 to 35 with $df = 12$. The negatively skewed curve (right) is the same as the positively skewed curve except that the order of levels is reversed. The resulting profiles are normalized such that the mean square of each curve becomes unity.

description). In the four-way analyses the mode reduction procedure was also used for the same reason.

3.1.   Three-way synthetic shifted data

Two three-way error-free shifted data sets were created following (5) in order to test the quasi-ALS procedure for the three-way SFA model. The profiles given in Figure 1 (type I) are the sequential factors used to generate the three-way shifted data. They are similar to those used for the two-way shifted data (Figure 2 in Part II [5]) but are even less steep than before and hence more difficult to recover. As before, 5 levels at each end are extra ones, included to fill the levels revealed after a maximum 5-level shift; this left 60 'within window' levels for mode A. True mode B and C factor loadings were sampled randomly from a uniform distribution of real numbers bounded by 0 and 1 exclusively. Unlike the two-way case, the numbers of levels of modes B and C were fixed at 15 and 10 respectively. Two kinds of shift values were used. Real numbers were used for one set of true shifts so that the resulting synthetic data better resembled real shifted data. These true shifts were randomly sampled from a uniform distribution of real numbers bounded by -5.5 and 5.5 exclusively. Another set of the true shifts was then obtained by rounding the sampled fractional shifts to

integers. The true **A**, **B**, **C** and **S** were then combined following (5) to produce two sets of three-way error-free shifted data. Fallible data were obtained by adding normally distributed error to the error-free shifted data sets. The error proportion in the fallible data is, on average, 15.1% of the total mean square and 36.4% of the total variance. The difference in the error proportion reflects the fact that the mean of the true part is 1.47 standard units while the expected mean in the error part is zero.

SFA, Parafac2 and Parafac1 were directly fit to each of the four $60 \times 15 \times 10$ shifted data sets, using 10 different random starts. We refer the reader to Reference [7] for a description of the direct fit Parafac2 algorithm used. SFA stopping criterion was the same as in the two-way weighted factor case: a maximum parameter change of 0.001% or a maximum of 1000 iterations in each of EIS1, EIS2 and FLS. For Parafac1 and Parafac2 the iteration limit was increased to 5000 but the parameter change criterion remained the same. Size weights (**A**, **B** and **C**) were constrained to be non-negative by using FNNLS [16], except for **A** in the direct fit Parafac2. Table I summarizes the recovery correlations and model fits of the best solution for each condition. The recovery correlations are averaged over the three factors within the solution.
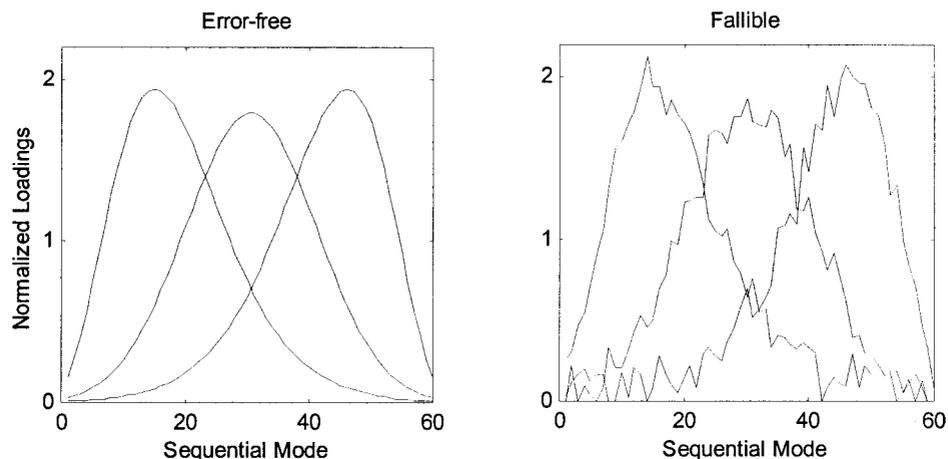
Figure 2. Recovered sequential factors from error-free (left) and fallible (right) three-way shifted data with integer shifts. Loadings are normalized such that each factor has unit mean square.

Table I. Recovery correlations and model fits (R^2) for three-way shifted data (type I)

		Integer shifts		Fractional shifts	
		Error-free	Fallible[a]	Error-free	Fallible[b]
Parameter recovery (r)					
SFA	A	1.0000	0.9883	0.9991	0.9878
	B	1.0000	0.9906	0.9999	0.9916
	C	1.0000	0.9929	1.0000	0.9920
	S	0.9194	0.7688	0.9217	0.8134
	S*B	1.0000	0.9154	0.9957	0.9201
PF2	B	0.9977	0.9577	0.9979	0.9575
	C	0.9949	0.9733	0.9953	0.9733
PF1	A	0.9761	0.9625	0.9781	0.9643
	B	0.9557	0.9489	0.9526	0.9476
	C	0.9967	0.9881	0.9969	0.9885
Fit to data (R^2)					
SFA		1.0000	0.6489	0.9999	0.6486
PF2		0.9988	0.7394	0.9988	0.7390
PF1		0.9670	0.6240	0.9679	0.6240

[a]Error proportion is 15.08% of total mean square and 36.36% of total variance; mean of the true part is 1.4709 standard units.
[b]Error proportion is 15.08% of total mean square and 36.38% of total variance; mean of the true part is 1.4722 standard units.

The quasi-ALS procedure perfectly recovers the factor weights from the error-free data for the integer shift case and nearly perfectly for the fractional shift case, confirming that latent trilinear factor structure that is disturbed by factor shifts can be recovered from three-way data by fitting the SFA model (5). It also suggests that, at least under some conditions, the factor loadings are uniquely determined and that the quasi-ALS algorithm can find them. The shift size parameters, however, are not as easy to recover. As noted in the two-way case, this is due to a few missed shift estimates. The recovery of weighted shifts **S*B** is perfect to four decimals for the integer shift case and near perfect for the fractional shift case.

For the fallible cases the recovery of parameters by SFA is better than expected. The recovery correlation is, most times, greater than 0.98 for **A**, **B** and **C** and greater than 0.90 for **S*B**, despite the fact that more than one-third of the total data variance is due to error.

When using the two types of true shifts, we initially expected to find an advantage of the fractional shift search over the integer search only when the true shifts were themselves fractional. However, there is no recognizable difference in fit and recovery between the integer and fractional shift data cases. FLS significantly improves fit and recovery for both cases, but it is hard to find a substantial difference between the two cases, either before or after FLS. There may be two reasons for this. First, the interpolation of fractional shifts may be no more advantageous for the fractional shift case than the integer case, because the sequential factors are so smooth that the two types of resulting shifted data are not substantially different. Second, FLS always improves fit and recovery even when true shifts are integers and the data are error-free, unless EIS perfectly recovers the integer shifts. Often FLS improves the shift estimates by 1 unit toward the true values, and this tends

to happen more frequently when the true sequential factors are somewhat difficult to recover (e.g. owing to smooth slowly changing profiles).

Figure 2 shows the recovered sequential factor profiles from the error-free and fallible shifted data sets. The right plot shows that SFA recovers the overall pattern of the sequential factors well, even when the amount of error is substantial.

It is surprising to find that Parafac2 (PF2 in Table I) also recovers **B** and **C** well; the recovery correlation is greater than 0.99 in the error-free case and greater than 0.95 in the fallible case. Even Parafac1 (PF1) recovers all factor loadings unexpectedly well, with a recovery correlation greater than 0.94 for all modes. This surprisingly good recovery by the Parafac models may have occurred because the sequential profiles given in Figure 1 are flatter and smooth rather than steep and/or jagged. As a result, the applied shifts are not sufficient to invalidate either the model requirement of Parafac1 that the position of the sequential factors is invariant for all levels of mode B, or that of Parafac2 that the angles between the sequential factors are invariant for all levels of mode B. That is, a maximum shift of 5 (or 5.5 in the fractional case) does not have as much effect on the values of these profiles as it would have, for example, with steep or jagged ones; thus even Parafac1 can recover the factors well.

To better understand the degree to which these data met the Parafac2 requirement of constant factor angles, we computed the cosines (i.e. congruence coefficients) between shifted profiles. Table II shows how much the angles between shifted factors (i.e. off-diagonal values in Φ) vary across the 15 mode B levels in the error-free integer-shifted case. We use cosines instead of correlations because all the profiles used in this study (and in spectral data in general) are positive and hence correlations, which are based on centered values, might be misleading. 'True' is the cosine between the two subscripted type I factors before they are shifted. The mean and standard deviation are computed over the 15 sets of cosines in which the variation is due to the different shifts. As expected, the shifting does not change the angles between the sequential factors much. Thus it is not surprising that Parafac2 recovers the factor loadings well.

We generated other sets of shifted data by using the profiles illustrated in Figure 3 in order to investigate the effect of violations of the constant-angle requirement and the constant-position requirement on the behavior of Parafac1, Parafac2 and SFA when the factor profiles are not so smooth. To save data-fitting time, the data size was reduced from $60 \times 15 \times 10$ to $30 \times 10 \times 7$ and the number of factors from three to two. All profiles in Figure 3 are identical except for their relative position. The two profiles are separated by 17 levels in type II and by 4 levels in type III. A set of integer shifts was chosen such that the factor angle changed

Table II. Variation of cosines in Φ due to independent shifting of the type I profiles

	True	Mean	SD
ϕ_{12}	0.6482	0.6503	0.1078
ϕ_{13}	0.6482	0.6479	0.1191
ϕ_{23}	0.2194	0.2296	0.0707
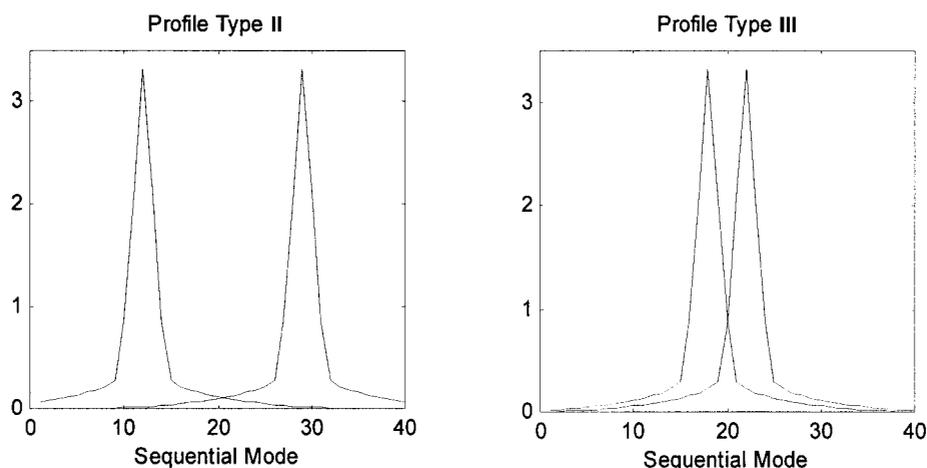
Figure 3. Sequential factors (types II and III). Forty levels were used to produce 30 window levels for the sequential mode in the resulting shifted data.

substantially across the 10 mode B levels for the type III profiles but not for the type II, again given a maximum shift of 5. Table III reports the true shifts and the resulting cosines for the two profile types. True mode B and C factor weights were randomly sampled from a uniform distribution of real values bounded by 0 and 1 exclusively. Since the two peaks are far from each other in type II, the shifting does not cause the cosine between them to increase substantially (its standard deviation is 0.026). In contrast, the shifting changes the cosine substantially in type III (its standard deviation is 0.316). Consequently, Parafac2 should be able to almost perfectly fit the resulting type II data, but not type III. The same convergence criterion and iteration limit were used for all model fitting as before.

Table IV summarizes the recovery correlations and model fits of the best solution out of five random starts for the error-free shifted type II and III data. SFA perfectly recovers the parameters for both data types. As expected, Parafac2 almost perfectly recovers **B** and **C** for the type II data but clearly fails for type III. The cosine recovered for the type III data by Parafac2 indicates that the solution is degenerate (in the precise mathematical sense [17–20]), and the degeneracy happened consistently. This consistent occurrence of degeneracy seems to be caused by the violation of the Parafac/Parafac2 requirement of invariant factor angle [17–20]. The model fails when the factor angle varies substantially across the levels of the shifting mode. Table IV also clearly shows that Parafac1 fails for both types of data, as we expected.

Both types of profiles are sufficiently steep that, given the maximum shift of 5, the Parafac1 requirement that the position of the sequential factors is invariant for all levels of the shifting mode does not hold.

3.2. Chromatographic data with retention time shifts

As mentioned earlier, Bro *et al.* [2] suggested using Parafac2 to resolve the position shift problem. They directly fit Parafac2 to a 28 (elution time) × 15 (sample) × 20 (excitation wavelength) × 78 (emission wavelength) chromatographic data set in which the elution profiles were assumed to shift differently from one sample to another. See Reference [2] for further details. We reanalyzed the same data to assess whether SFA and Parafac2 are appropriate. We used an edited version: first every other level of the original 78 emission levels (250–560 nm emission with a 4 nm interval) was taken, and then the first two resulting levels (250 and 258 nm), in which all values are missing, were deleted, resulting in a 28 × 15 × 20 × 37 four-way chromatographic data array.

Table III. Shifts and resulting cosines (ϕ_{12}) of type II and III profiles

Mode B level	Factor 1	Factor 2	ϕ_{12} of type II	ϕ_{12} of type III
1	−3	−5	0.0115	0.5551
2	−1	0	0.0260	0.1720
3	4	1	0.0070	0.8470
4	−1	−3	0.0115	0.5551
5	−5	3	0.0969	0.0595
6	−2	−1	0.0260	0.1720
7	5	2	0.0070	0.8470
8	3	0	0.0070	0.8470
9	1	3	0.0322	0.1479
10	1	−1	0.0115	0.5551

Table IV. Recovery correlations and model fits (R^2) for three-way shifted data (types II and III)

		Type II	Type III
Parameter recovery (r)			
SFA	**A**	1.0000	1.0000
	B	1.0000	1.0000
	C	1.0000	1.0000
	S	1.0000	1.0000
	S∗B	1.0000	1.0000
PF2	**B**	0.9996	0.8491
	C	0.9997	0.6888
	ϕ_{12}	0.0936	−0.9894
PF1	**A**	0.4897	0.7348
	B	0.4749	0.4329
	C	0.5210	0.5093
Fit to data (R^2)			
SFA		0.9999	1.0000
PF2		0.9999	0.9949
PF1		0.4515	0.6892

By taking advantage of the fact that multilinearity holds well in the excitation and emission modes but not in the elution and sample modes, because the elution profiles shift differently across samples, Bro *et al.* [2] devised an elegant check of the data structure for this unusual four-way case. Since the shift problem is supposed to be independent of both the excitation and emission modes, and the original data set is four-way, it is possible to obtain an ideal reference solution against which a competing model can be assessed. The $28 \times 15 \times 20 \times 37$ four-way data were first unfolded into a $420 \times 20 \times 37$ three-way array by merging the elution and sample modes into what Tucker [21] and Kroonenberg [22] call a 'combination mode'. Since the 420 levels in the combination mode can then be considered as independent measurement units, it is reasonable to expect trilinearity in the unfolded data so far as multilinearity holds in the excitation and emission modes.

Four-way Parafac2 and four-way SFA are considered here as two competing models for fitting these chromatographic data. The four-way Parafac2 model has the form

$$\mathbf{X}'_{jk}\mathbf{X}_{jk} = \mathbf{D}\langle\mathbf{c}_k\rangle\langle\mathbf{b}_j\rangle\mathbf{\Phi}\langle\mathbf{b}_j\rangle\langle\mathbf{c}_k\rangle\mathbf{D}' + \mathbf{E}_{jk}$$
$$\mathbf{\Phi} = \mathbf{A}'_{jk}\mathbf{A}_{jk} \tag{22}$$

where $\mathbf{X}'_{jk}\mathbf{X}_{jk}$ is an $L \times L$ (L being the number of levels in mode D) cross-product matrix at level j in mode B and level k in mode C. It implies that the relationships (cosines) between factors in mode A are invariant across all combinations of the mode B and C levels, although $\mathbf{A}$ itself is free to vary, even in its number of levels, across these combined levels. We can fit (22) to the data by taking the elution mode as mode A and the sample mode as mode B. The Parafac2 fitting implies that the shifting along the elution mode does not change the interrelation of the sequential factors across all combinations of a sample and an excitation (or emission) wavelength. The invariance of factor angles must hold across the excitation or emission wavelengths for multilinearity to hold in these modes.

The four-way SFA model can easily be drawn from the *n*-way SFA model (6) by taking the shifted mode (mode A), the shifting mode (mode B) and two other modes (modes C and D) to be related as

$$\mathbf{X}_{jl} = \oslash f_j(\mathbf{A})\langle\mathbf{b}_j\rangle\langle\mathbf{d}_l\rangle\mathbf{C}' + \mathbf{E}_{jl} \tag{23}$$

where $\mathbf{X}_{jl}$ is the $I \times K$ data matrix at level j and l in the $I \times J \times K \times L$ four-way data array $\underline{\mathbf{X}}$. When modeling the chromatographic data using four-way SFA, it is preferable that the shifting pattern of the elution profiles is independent from one sample to the next because this provides additional information to help uniquely identify the factors.

The number of factors was set to four, following the result of the dimensionality test performed in Reference [2]. There were many missing cells in the fully crossed four-way chromatographic data array (36.9%), since it is impossible to obtain information on the emission levels below excitation. To handle the missing elements, an alternating weighted least squares procedure was used in the Parafac1 fitting. Weights of zero were assigned to both the missing data cells and those combinations of fixed parameters in the

ALS estimation that corresponded to the missing elements in the data. Iterative imputation adopted for the four-way Parafac2 and SFA replaced missing elements with the values predicted by the model after every iteration, which also gives a least squares solution [23].

The same stopping criterion and iteration limits were used for Parafac1 and Parafac2 as in the synthetic data analysis. However, because the quasi-ALS procedure for the four-way SFA takes much longer than the standard ALS estimation, even with two of the time-saving strategies noted in the Appendix of Part II [5] (two-stage EIS and periodic shift search), the stopping criteria had to be relaxed: a maximum parameter change of 0.01% and an iteration limit of 500 in each stage of the SFA. Since the factor weights correspond to amounts of physical entities in all modes, factor loading parameters were constrained to be non-negative in all analyses except for the elution mode in the Parafac2 analysis. The FNNLS algorithm [16] was used again for the non-negativity constraint to save time. Ten different random starts were tried in both the three-way Parafac1 and four-way Parafac2 analyses. However, only two were applied for the four-way SFA, since it takes much longer, particularly when there are as many as four factors*.

Figure 4 shows the spectra in the excitation and emission modes estimated by three-way Parafac1, and four-way Parafac2 and SFA. It is clear in Figure 4 that Parafac2 successfully resolved the position-shifting problem in the chromatographic data. The factor loadings correlate highly between the Parafac1 and Parafac2 solutions, ranging from 0.9898 to 0.9996 in the excitation mode and from 0.9603 to 0.9999 in the emission mode. SFA, however, recovers a pair of almost perfectly redundant factors in the excitation and emission modes; factors 2 and 4 have a correlation of 0.9984 in the excitation mode and 0.9959 in the emission mode. These factors correlate highly with the Parafac1 factor 2 (0.9968 in the excitation mode and 0.9980 in the emission mode). The SFA factor 1 seems to maximally approximate both Parafac1 factors 1 and 4, while the remaining factor (F3) in these two solutions also matches well ($r = 0.9671$ in the excitation mode and 0.9619 in the emission mode). Thus SFA is only partly successful in recovering the latent structure of the chromatographic data. Table V summarizes the model fits of the reference (Parafac1), Parafac2 and SFA solutions.

We further investigated the Parafac1 reference solution to better understand the latent structure of the chromatographic data. First we reconstructed separate elution and sample modes for each of the 10 Parafac1 solutions by reshaping each 420×4 estimated factor loading matrix of the combination mode into a $28 \times 15 \times 4$ parameter array. (Recall that these parameters were estimated not by a fully crossed multilinear model but by the less restrictive model that does not require multilinearity between the elution and sample modes.) The size of shifts for each factor was then obtained from this reshaped parameter array by taking the relative difference in the highest peak location of elution profiles among the 15 samples. The resulting shift matrices were highly reliable (identical in nine out of the 10 solutions).

*In the time since this study was conducted, desktop computer speeds have increased by a factor of 10–20, making SFA of larger problems much more practical.

Excitation   Emission
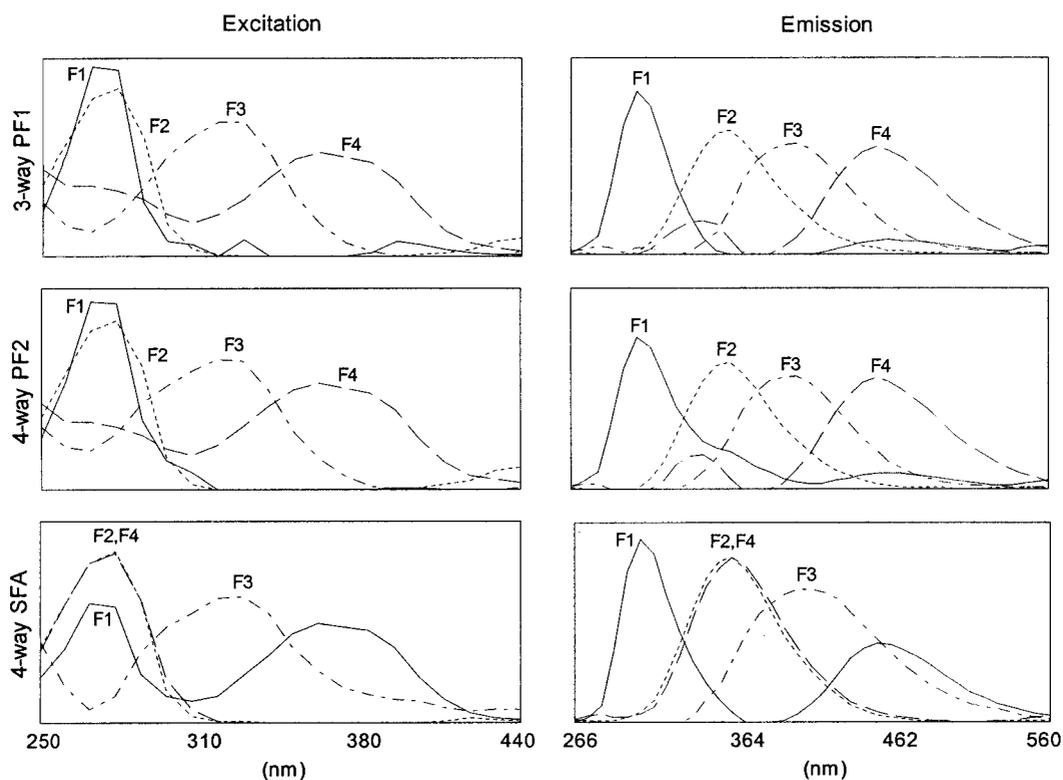
Figure 4. Excitation (left) and emission (right) spectra estimated by three-way Parafac1 (top), four-way Parafac2 (middle) and four-way SFA (bottom). All factors are numbered so that they are comparable across the three models and consistent between the excitation and emission modes.

Figure 5 shows that the shifting pattern of elution profiles is fairly similar across factors in most samples, which implies that, for most of the chromatographic data, shifting can be interpreted in terms of shifts at the 'surface' or data level instead of the latent level.

In addition to the position shifting, the elution profiles in the Parafac1 reference solution were found to change their shape substantially for some samples, although in theory they should be identical [2]. Therefore we edited the chromatographic data to use only those samples for which the

shape of the elution profiles is relatively homogeneous. Selected are samples 3, 4, 5, 6, 7, 9 and 15, providing a $28 \times 7 \times 20 \times 37$ four-way data array. Then we fit the four-way SFA to the resulting 'more homogeneous' chromatographic data to get a possible indication of whether the shape change of elution profiles had been an important cause of the redundant factors. However, the redundant factors did not disappear. Both data sets seem to have a very similar SFA factor structure; the factor loading correlations range from 0.9740 to 0.9997 in the excitation mode and from 0.9917 to 0.9998 in the emission mode.

The redundant factors could have occurred because the solutions were locally optimal (or simply prematurely stopped). Although both more random starts and a more stringent convergence criterion are required to increase the chance of obtaining the global minimum, it was not practical to apply them at the time these studies were done because of the heavy computational load of the quasi-ALS procedure

Table V. Model fits by Parafac1, Parafac2 and SFA

Model	MSE	R^2
Three-way PF1	154.2466	0.9893
Four-way PF2	216.4249	0.9850
Four-way SFA	479.2075	0.9474
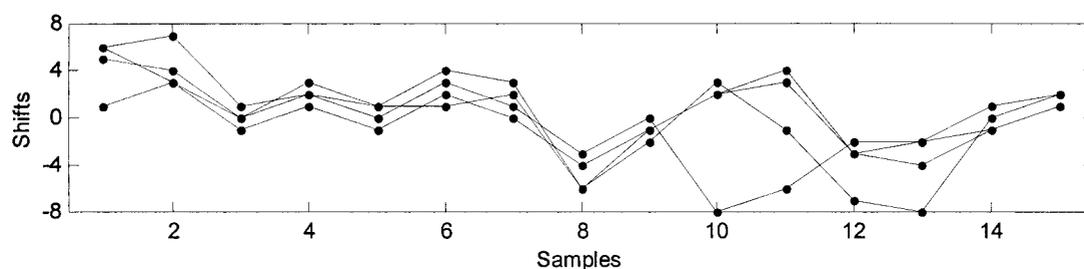
Figure 5. Shifts inferred from the three-way Parafac1 solution, rounded to integers, and centered for each factor.
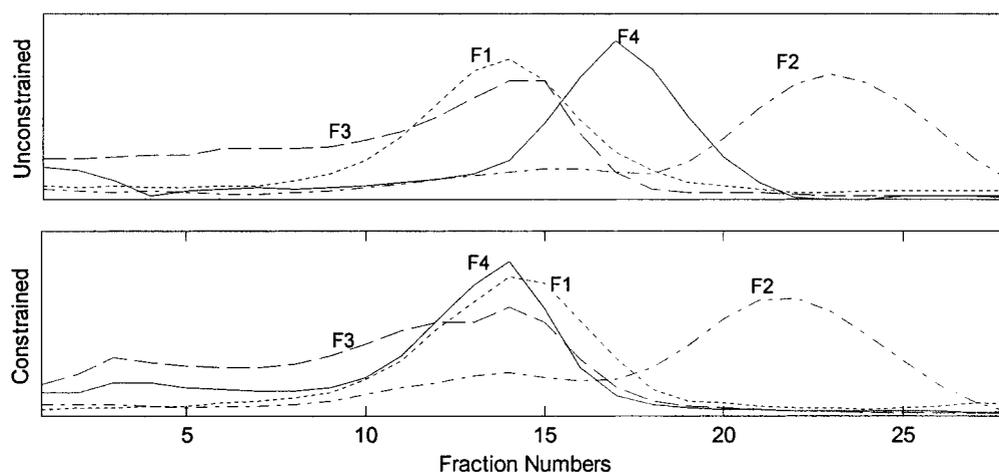
Figure 6. Elution profiles from unconstrained (top) and constrained (bottom) SFA. Factors are numbered to be consistent with the SFA factors in Figure 4.

for the four-way data. Instead we tried a different remedy. Given the 'reference' shifts inferred from the Parafac1 solution and the Parafac1 excitation and emission factor weights, it is possible to obtain an ideal solution of the factor weights for the elution and sample modes. By holding **S**, **C** and **D** fixed at the reference solution values and fitting four-way SFA, we obtained the ideal solution for **A** and **B**. The elution profiles estimated with the constrained values are compared with those from the unconstrained SFA in Figure 6. The profiles match well for all but factor 4, which is positioned farther from factor 2 in the constrained solution than in the unconstrained one.

In order to understand why the redundant factors occur, we again consider the spectra estimated by different models in Figure 4. As noted previously, the SFA factors 2 and 4 are almost identical and the SFA factor 1 approximates both Parafac1 factors 4 and 1. To further investigate which SFA factor corresponds to which Parafac1 factor, we compared the factor sizes among models. Table VI shows each factor's percentage of the sum of squared factor weights in the mode that contains the original scale of the data, summarized by models. The sum of squared factor weights is different from the variance accounted for by factors, since the fit data are not centered. However, it can be a good approximation of the relative magnitude of factor contribution to the data, in that all factors are theoretically non-negative and hence the additive constant in factor loadings might be more or less evenly spread out across factors. Given Figure 4 and Table VI, it seems to be the case that the Parafac1 factor 2 splits into the unconstrained SFA factors 2 and 4 and that the

Table VI. Percentages of sum of squared factor[a] weights in scaled mode

	F1	F2	F3	F4
Parafac1	11.3	66.3	9.4	13.0
Parafac2	14.8	61.0	10.9	13.4
Constrained SFA	10.8	69.1	8.4	11.6
Unconstrained SFA	28.9	44.8	16.8	9.5

[a]Factors are numbered to be consistent with those in Figures 4 and 6.

unconstrained SFA factor 1 fits the same portion that the Parafac1 factors 1 and 4 explain.

It is perhaps not surprising that the constrained SFA percentages are very close to those for Parafac1, since **S**, **C** and **D** are predicted by Parafac1. Another interesting observation in the constrained SFA solution is that the mean square error (or badness of fit) has increased substantially, compared with the Parafac1 solution, when the elution and sample modes are further decomposed. For example, MSE increases from 105.6462 (Parafac1) to 707.5414 (constrained SFA) when the 7-sample data are analyzed. Because Parafac1 provides **S**, **C** and **D** and the other two modes are not decomposed in either the three-way Parafac1 or the four-way Parafac2 analysis, we can attribute the large MSE increase to the 'badness' of the bilinearity between the elution and sample modes.

To further test whether the SFA solution was a local optimum, we initialized all parameters except shifts with the best estimates: the excitation and emission weights from the Parafac1 reference solution, and the elution and sample weights from the constrained SFA solution. However, this 'best' starting position also soon approached the 'redundant' factor solution in EIS1, after about 40–50 iterations.

To make more random starts feasible, we reconstructed two three-way chromatographic data sets (elution × sample × excitation and elution × sample × emission) based on the reference solution, using the mode reduction procedure. As noted in the Appendix of Part II [5] such reconstructed data are hypothetical and subject to the validity of the reference solution. We fit the three-way SFA model (5) to the reconstructed data sets by using 10 random starts and with a more stringent convergence criterion: a maximum parameter change of 0.001% and an iteration limit of 1000 in each stage. The redundant factors persistently appeared in all three-way SFA solutions. Thus we consider it unlikely that the occurrence of the redundant factors is simply due to a local optimum.

There are other possible reasons why the Parafac1 factor 2 is split up into the SFA factors 2 and 4 or, equivalently, why the SFA factor 1 fits both the Parafac1 factors 1 and 4. First, as already noted, the shifting pattern of the four factors in

Figure 5 is very similar in most samples. In the error-free case the lack of independence in shifting pattern would not impair the unique and valid recovery of parameters from the shifted data in which multilinearity is otherwise perfectly exhibited. In some fallible cases where error is fit, however, it might cause SFA to have problems in identifying the sequential factors. Second, SFA seems to compromise the excitation and emission weights in order to further decompose the elution and sample modes in which (bi)linearity is only partly present. Third, the elution profiles estimated by SFA look very similar to one another (e.g. constrained SFA factors 1, 2 and 4 in Figure 6). If the distance between the peaks of two such similar sequential factors is within the allowed shifting distance in the SFA fitting, they may be easily confused with each other during the estimation of their positions. We consider this a new type of indeterminacy in SFA. In conclusion, it is not clear which, if any, of the proposed explanations for the redundancy in the SFA solutions is correct, and so this issue is unresolved.

4. DISCUSSION

We have shown how the two-way shifted factor model [5] and its estimation algorithms can be easily extended to multiway (i.e. higher than two-way) cases, at least when the added mode(s) simply reweight the shifted factors. Standard multilinear models such as Parafac1 are already unique under certain conditions (e.g. the Kruskal conditions for factor independence [24]). Nonetheless, in cases where some sequential factors are independently shifted, SFA can take advantage of this added source of systematic latent variation to provide even more robust and accurate recovery of the latent structure.

As with the two-way SFA, the multiway shifted factor models have several characteristics not present in the simple multilinear case: (a) factors shifting independently (i.e. at the latent level) are easier to identify than those that shift simultaneously (i.e. at the data level); (b) shifted factors are easier to identify when the sequential factor profiles are distinctive from one another in shape than when they are similar; (c) steep and/or jagged sequential factors are easier to identify than flatter and/or smooth factors [5,10]; and (d) the mode B (i.e. shift-controlling mode) factor weights moderate not only factor influence, but also the impact of shifting.

When profiles shift along a sequential mode, simple fitting of Parafac1 is no longer appropriate. If the shifting has occurred at the data level (i.e. sequential factors are shifted simultaneously), it can be undone at the data level by preprocessing the shifted data so that the sequential data profiles are lined up. Then one can fit Parafac1 to the position-aligned sequential data as mentioned in Part I [4]. A simple example of such a preprocessing method is maximization of cross-correlations between sequential data profiles, as Cattell [25] did in his time-corrected P-technique. Most sophisticated data-preprocessing techniques for the position shift problem include, among others, time warping [26,27], gradient-based motion estimation [28,29], minimization of structural complexity due to shift variance [30,31], and regional shift alignment by partial linear fit [32].

As a means of dealing with factor shifts, the shifted factor analysis method has both advantages and disadvantages compared with Parafac2. SFA relaxes the multilinearity requirement of Parafac1 in mode A (i.e. the shifted mode) by allowing sequential factors to shift in position. Parafac2 also relaxes the multilinearity but in a different way. It allows sequential factors (i.e. the mode over which cross-products are computed) to be anything that fulfils the Parafac2 condition of invariant factor angles (i.e. Φ, for example in (22)) over the shifting mode. When sequential factors are independently shifted and invariant in shape, SFA properly resolves the non-multilinearity caused by the latent level shifting. When sequential factors change shape as well as position, Parafac2 can be more useful so long as the angles between sequential factors do not change too much across levels of mode B. When the invariant factor angle condition is seriously violated, the Parafac2 solution becomes consistently degenerate.

Another important distinction between SFA and Parafac2 is that sequential factors are explicitly modeled in SFA as a common set of position-shifted profiles which are the same at all levels of the other modes. In contrast, they are allowed in Parafac2 to change over the shifting mode in any way that preserves the angles between them. Thus, if it is theoretically important to identify a *common, invariant* set of sequential factors that are position-shifted independently over a data mode (i.e. the shift-controlling mode), SFA is the preferred approach.

The current quasi-ALS procedure becomes very time consuming as the data array and/or model dimensionality becomes large, particularly in the number of data modes. In our experience (with a 400 MHz processor) it is practical up to three-way cases such as the $60 \times 15 \times 10$ synthetic shifted data, with three factors and 11 allowed shifts in the analysis. For example, with the speed-up provisions, the two-stage EIS and periodic shift search every 10th iteration, one random start took about 30 min to converge. The increase in the computation time is approximately exponential to the number of factors and multiplicative to the maximum allowed shift and the number of levels in all modes. Even when the two-stage EIS and the periodic shift search procedures are used, the SFA fitting of the four-way chromatographic data was too slow to apply, say, 10 random starting positions. Thus faster computers and further algorithmic development will be needed for higher-than-three-way cases and large three-way cases.

REFERENCES

1. Field AS, Graupe D. Topographic component (parallel factor) analysis of multichannel evoked potentials: practical issues in trilinear spatiotemporal decomposition. *Brain Topogr*. 1991; **3**: 407–423.

2. Bro R, Anderson CA, Kiers HAL. PARAFAC2—Part II. Modeling chromatographic data with retention time shifts. *J. Chemometrics* 1999; **13**: 295–309.

3. Tauler R. Multivariate curve resolution applied to second order data. *Chemometrics Intell. Lab. Syst.* 1995; **30**: 133–146.

4. Harshman RA, Hong S, Lundy ME. Shifted factor analysis. Part I. Models and properties. *J. Chemometrics* 2003; **17**: 363–378.

5. Hong S, Harshman RA. Shifted factor analysis. Part II. Algorithms. *J. Chemometrics* 2003; **17**: 379–388.

6. Bro R. *Multi-way Analysis in the Food Industry: Models, Algorithms and Applications*. University of Amsterdam: Amsterdam, 1998. (http://www.models.kvl.dk/users/asmus/thesis/thesis.html).

7. Kiers HAL, ten Berge JMF, Bro R. PARAFAC2—Part I. A direct fitting algorithm for the PARAFAC2 model. *J. Chemometrics* 1999; **13**: 275–294.

8. Harshman RA. PARAFAC2: mathematical and technical notes. *UCLA Working Papers Phonet.* 1972; **22**: 30–44. (http://publish.uwo.ca/~harshman/wpppfac2.pdf).

9. Harshman RA. An index formalism that generalizes the capabilities of matrix notation and algebra to *n*-way arrays. *J. Chemometrics* 2001; **15**: 689–714.

10. Hong S. Shifted factor analysis: a test of models and algorithms. *Master's Thesis*, University of Western Ontario, London, ON, 1997.

11. Carroll JD, Chang JJ. Analysis of individual differences in multidimensional scaling via *N*-way generalization of 'Eckart–Young' decomposition. *Psychometrika* 1970; **35**: 283–319.

12. Harshman RA. Foundations of the PARAFAC procedure: models and conditions for an 'explanatory' multi-modal factor analysis. *UCLA Working Papers Phonet.* 1970; **16**: 1–84. (http://publish.uwo.ca/~harshman/wpppfac0.pdf).

13. Harshman RA, Hong S. 'Stretch' vs. 'slice' methods for representing three-way structure via matrix notation. *J. Chemometrics* 2002; **16**: 198–205.

14. Harshman RA. Shifted factor analysis models and estimation expressed in array notation. *Department of Psychology Res. Bull. #762*, University of Western Ontario, London, ON, 2001. (http://publish.uwo.ca/~harshman).

15. Khatri CG, Rao CR. Solutions to some functional equations and their applications to characterization of probability distributions. *Sankhya, Ser. A* 1968; **30**: 167–180.

16. Bro R, de Jong S. A fast non-negativity-constrained linear least squares algorithm for use in multi-way algorithms. *J. Chemometrics* 1997; **11**: 393–401.

17. Kruskal JB, Harshman RA, Lundy ME. How 3-MFA data can cause degenerate PARAFAC solutions, among other relationships. In *Multiway Data Analysis*, Coppi R, Bolasco S (eds). Elsevier: Amsterdam, 1989; 115–122.

18. Lundy ME, Harshman RA, Kruskal JB. A two-stage procedure incorporating good features of both trilinear and quadrilinear models. In *Multiway Data Analysis*, Coppi R, Bolasco S (eds). Elsevier: Amsterdam, 1989; 123–130. (http://publish.uwo.ca/~harshman/pfcore89.pdf).

19. Paatero P. Construction and analysis of degenerate PARAFAC models. *J. Chemometrics* 2000; **14**: 285–299.

20. Mitchell BC, Burdick DS. Slowly converging PARAFAC sequences: swamps and two-factor degeneracies. *J. Chemometrics* 1994; **8**: 155–168.

21. Tucker LR. Some mathematical notes on three-mode factor analysis. *Psychometrika* 1966; **31**: 279–311.

22. Kroonenberg PM. *Three-mode Principal Component Analysis: Theory and Applications*. DSWO Press: Leiden, 1983.

23. Kiers HAL. Weighted least squares fitting using ordinary least squares algorithms. *Psychometrika* 1997; **62**: 251–266.

24. Kruskal JB. Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra Appl.* 1977; **18**: 95–138.

25. Cattell RB. The structuring of change by P-technique and incremental R-technique. In *Problems in Measuring Change*, Harris CW (ed.). University of Wisconsin Press: Madison, WI, 1963; 167–198.

26. Nielsen N-PV, Carstensen JM, Smedsgaard J. Aligning of single and multiple wavelength chromatographic profiles for chemometric data analysis using correlation optimized warping. *J. Chromatogr. A* 1998; **805**: 17–35.

27. Bylund D, Danielsson R, Malmquist G, Markides KE. Chromatographic alignment by warping and dynamic programming as a pre-processing tool for PARAFAC modeling of liquid chromatography–mass spectrometry data. *J. Chromatogr. A* 2002; **961**: 237–244.

28. Westad F, Martens H. Shift and intensity modeling in spectroscopy—general concept and applications. *Chemometrics Intell. Lab. Syst.* 1999; **45**: 361–370.

29. Horn BKP, Schunck BG. Determining optical flow. *Artific. Intell.* 1981; **17**: 185–203.

30. Prazen BJ, Synovec RE, Kowalski BR. Standardization of second-order chromatographic/spectroscopic data for optimum chemical analysis. *Anal. Chem.* 1998; **70**: 218–225.

31. Fraga CG, Prazen BJ, Synovec RE. Comprehensive two-dimensional gas chromatography and chemometrics for the high-speed quantitative analysis of aromatic isomers in a jet fuel using the standard addition method and an objective retention time alignment algorithm. *Anal. Chem.* 2000; **72**: 4154–4162.

32. Vogels JTWE, Tas AC, Venekamp J, Van Der Greef J. Partial linear fit: a new NMR spectroscopy preprocessing tool for pattern recognition applications. *J. Chemometrics* 1996; **10**: 425–438.