British Journal of Mathematical and Statistical Psychology (2009), 62, 601–620
© 2009 The British Psychological Society



The British Psychological Society

www.bpsjournals.co.uk

Discriminating between strong and weak structures in three-mode principal component analysis

Eva Ceulemans[1]* and Henk A. L. Kiers[2]

[1]Katholieke Universiteit Leuven, Department of Educational Sciences, Leuven, Belgium

[2]Rijksuniversiteit Groningen, Groningen, The Netherlands

Recently, a number of model selection heuristics (i.e. DIFFIT, CORCONDIA, the numerical convex hull based heuristic) have been proposed for choosing among Parafac and/or Tucker3 solutions of different complexity for a given three-way three-mode data set. Such heuristics are often validated by means of extensive simulation studies. However, these simulation studies are unrealistic in that it is assumed that the variance in real three-way data can be split into two parts: structural variance, due to a true underlying Parafac or Tucker3 model of low complexity, and random noise. In this paper, we start from the much more reasonable assumption that the variance in any real three-way data set is due to three different sources: (1) a strong Parafac or Tucker3 structure of low complexity, accounting for a considerable amount of variance, (2) a weak Tucker3 structure, capturing less prominent data aspects, and (3) random noise. As such, Parafac and Tucker3 simulation studies are run in which the data are generated by adding a weak Tucker3 structure to a strong Parafac or Tucker3 one and perturbing the resulting data with random noise. The design of these studies is based on the reanalysis of real data sets. In these studies, the performance of the numerical convex hull based model selection method is evaluated with respect to its capability of discriminating strong from weak underlying structures. The results show that in about two-thirds of the simulated cases, the hull heuristic yields a model of the same complexity as the strong underlying structure and thus succeeds in disentangling strong and weak underlying structures. In the vast majority of the remaining third, this heuristic selects a solution that combines the strong structure and (part of) the weak structure.

1. Introduction

Three-way three-mode data occur in many research domains. For example, one may think of data that indicate the extent to which a number of individuals display a number of behaviours in a number of situations. Data on the performance of a number of children on a set of mathematics tests at a number of time points constitute a

*Correspondence should be addressed to Dr Eva Ceulemans, Department of Educational Sciences, Andreas Vesaliusstraat 2, B-3000 Leuven, Belgium (e-mail: Eva.Ceulemans@ped.kuleuven.be).

DOI:10.1348/000711008X369474

second example. To gain insight into the often complex structure of such data, various exploratory three-mode data analysis techniques are available. The two most popular three-mode methods, Parafac (Carroll & Chang, 1970; Harshman, 1970) and Tucker3 analysis (Kroonenberg & De Leeuw, 1980; Kroonenberg, 1983; Tucker, 1966), belong to the family of three-mode principal component techniques. Being three-mode generalizations of standard principal component analysis (PCA), Parafac and Tucker3 summarize the main information in a three-way three-mode data set by reducing the three modes of the data to a few components and defining a linking structure between the three sets of components. Formally, a Parafac and Tucker3 analysis of an $I \times J \times K$ data array $\underline{\mathbf{X}}$ boils down to fitting the model

$$x_{ijk} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} a_{ip} b_{jq} c_{kr} g_{pqr} + e_{ijk}, \tag{1}$$

where P, Q, and R denote the number of components to which the three modes of the data are reduced, a_{ip}, b_{jq} and c_{kr} denote the entries of the component matrices **A**, **B**, and **C** of orders $I \times P$, $J \times Q$, and $K \times R$, g_{pqr} denotes the entries of the $P \times Q \times R$ core array $\underline{\mathbf{G}}$, which defines a linking structure between the three sets of components, and e_{ijk} denotes the error term or residual for data entry x_{ijk}. The difference between Parafac and Tucker3 is that in Parafac (i) the three modes are reduced to the same number of components (i.e. $P = Q = R$) and (ii) the core array is constrained to a superidentity array (i.e. $g_{pqr} = 1$ if $p = q = r$, and $g_{pqr} = 0$ otherwise), implying that the linking structure among the three sets of components takes the form of a one-to-one correspondence. As such, in the Parafac case, model formula (1) reduces to

$$x_{ijk} = \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr} + e_{ijk}. \tag{2}$$

As is the case for most exploratory data analysis techniques, an important challenge when applying Parafac and Tucker3 to real data sets is solving the model selection problem, that is, deciding which values of P, Q, and R, yield the most useful description of the data; note that some authors refer to this particular model selection problem as the dimensionality or complexity selection problem (see Kroonenberg, 2008). To help researchers with this often difficult task, a number of model selection heuristics have recently been proposed. First of all, Timmerman and Kiers (2000) introduced the DIFFIT method, which is a Tucker3 model selection tool. Additionally, building on earlier work by Bro (1998), Bro and Kiers (2003) presented the CORCONDIA method for choosing among Parafac solutions of different complexities. Finally, formalizing the visual inspection based model selection method proposed by Kroonenberg and Van der Voort (1987), Kroonenberg and Oort (2003), and Murakami and Kroonenberg (2003), Ceulemans and Kiers (2006) proposed the numerical convex hull based model selection heuristic, which can be used to select among three-mode principal component solutions of different types (e.g. Tucker3 and Parafac) and complexities (i.e. different values of P, Q and R).

To evaluate the performance of these model selection heuristics, different strategies have been followed. Bro and Kiers (2003) mainly demonstrated the usefulness of CORCONDIA by reanalysing real data sets, showing that the CORCONDIA results corroborated the conclusions from prior analyses. In contrast, DIFFIT and the numerical convex hull based model selection heuristic were validated by applying these techniques to a large number of simulated data sets, which were generated by

perturbing a three-mode principal component structure of low complexity (i.e. small *P*, *Q*, *R*) with random noise (Ceulemans & Kiers, 2006; Kiers & der Kinderen, 2003; Timmerman & Kiers, 2000). These simulation studies showed that, for such data, the performance of DIFFIT and the hull heuristic is nearly perfect, with the hull heuristic slightly outperforming DIFFIT.

In this paper, we will further refine the latter evaluation strategy, as it can be argued that the underlying assumption – that the variance in real three-way data can be split into two parts: structural variance, due to a true underlying three-mode principal component model of low complexity, and random noise – is unrealistic. For instance, for the simple, case of standard PCA, it is much more reasonable to assume that any two-way two-mode data set contains, apart from random noise, two sets of components: (i) a few 'strong' components, which each explain a considerable amount of variance in the data and thus describe the most salient aspects of the data; and (ii) a number of 'weak' components, which each capture a rather small amount of (co)variance and thus summarize more subtle or isolated data aspects. In a similar vein, Tucker, Koopman and Linn (1969) stated for the closely related technique of factor analysis that possible influences on the data can be divided into three categories: (i) a major factor domain representing the area of phenomena of interest to the experimenter; (ii) a minor factor domain representing other influences outside the area of interest; and (iii) specific and measurement error influences on the results of each variable. Generalizing to the three-mode case, we argue that the variance in any three-mode data set is due to three different sources: (i) a strong three-mode principal component structure of low complexity, accounting for the most salient relations in the data; (ii) a weak three-mode principal component structure, capturing less prominent relations in the data; and (iii) random noise.

Given this distinction between major and minor factor domains or strong and weak component structures, many applications researchers will want to focus on the major factors or strong component structures, as these represent the most prominent relations in the data. After all, it is virtually impossible to study all relations that are present in a data set, even though the weakness of some relations might be due to undersampling. To discriminate between these major and minor factors or strong and weak structures, researchers often rely on model selection techniques. For factor analysis, Tucker *et al.* (1969) investigated the extent to which a number of model selection heuristics – such as Guttman's (1954) lower and upper bounds for the number of common factors – are effective indicators for the number of factors in the major factor domain. However, for three-mode component analysis, given the above-mentioned set-up of the DIFFIT and hull simulation studies and the lack of an extensive CORCONDIA simulation study, we have no indication whether these techniques indeed succeed in disentangling strong and weak structures. Therefore, in this paper, we will run Parafac and Tucker3 model selection simulations in which the data are generated by adding a weak three-mode principal component structure to a strong one (Parafac and Tucker3, respectively), and perturbing the resulting data with random noise. To come up with a realistic design for this simulation study – how to divide the structural variance over the strong and weak structures – we start by analysing the residuals of real data sets, for which there is a reasonable consensus about the complexity of the strong underlying three-mode principal component model. In these simulations, we will primarily focus on the performance of the hull heuristic, as it is the only heuristic which was explicitly designed for choosing among Tucker3 as well as Parafac solutions.

An additional aim of the present paper is to empirically compare CORCONDIA and the hull heuristic. For this purpose, on data sets with a strong Parafac structure we will

also apply CORCONDIA. This comparison is of interest for two reasons. First, as mentioned above, no extensive CORCONDIA simulation study, manipulating several independent variables, has been reported up till now. Second, the rationale behind the hull heuristic and CORCONDIA is completely different. In particular, the hull heuristic is a more traditional model selection tool in that it selects the solution with the best balance of complexity and fit to the data. In contrast, CORCONDIA is tailor-made for the Parafac model selection problem, choosing the Parafac solution of the highest complexity that is still appropriate, that is, for which removing the restriction that the core array has to take the form of a superidentity array hardly changes its values. The details of both heuristics will be presented in Section 3.

The remainder of this paper is organized as follows. In Section 2, we analyse the residuals of six thoroughly analysed data sets, to empirically validate our claim that real three-way three-mode data contain, next to a strong underlying three-mode principal component structure, also a weak three-mode principal component structure, and to come up with a realistic design for our simulation study. In Section 3, the numerical convex hull based model selection heuristic and the CORCONDIA method are recalled. In Section 4, simulation studies are reported for data with strong Tucker3 structures (Tucker3 simulation study) and strong Parafac structures (Parafac simulation study), respectively, in which we evaluate to what extent the hull heuristic succeeds in discriminating between strong and weak three-mode principal component structures. Furthermore, in the Parafac simulation study, we also evaluate the performance of CORCONDIA. Section 5 discusses our results.

2. Analysis of real data sets

In this section, we will reanalyse six real three-way three-mode data sets, for which there is a reasonable consensus about the complexity of the strong underlying three-mode principal component model: the Osgood data, the anxiety data, the Chopin data, the bread data, the tongue data, and the cars and stars data. In particular, we will obtain weak three-mode principal component structures for these data sets by further analysing the residuals (i.e. the e_{ijk} in model formula (1)), once the strong underlying three-mode principal component solutions have been fitted to these data. The aim of these analyses is twofold. First, we wish to provide empirical support for our claim that the assumption behind the Ceulemans and Kiers (2006) simulation study on the performance of the hull heuristic – that the residuals of a three-mode principal component analysis can be considered random noise – is unrealistic. We will do so by demonstrating that a three-mode principal component structure of low complexity explains considerably more variance of the residuals e_{ijk} of real data than of randomly sampled data sets of the same size. Second, to ensure that the design of the current simulation studies is realistic, we wish to get a general idea of how the variance of real data sets is divided over strong and weak three-mode principal component structures of different complexities.

In the following paragraphs, we will first discuss the details of our analyses. Subsequently, the results with respect to the two aims described above will be presented.

2.1. Details of our analyses

The six real data sets were pre-processed and analysed as described in the reference articles listed in Table 1. This table further provides details with respect to the size of the

data, the type and complexity of the original analysis, and the resulting fit percentage, which we will denote by f_{strong} as it is a measure of the amount of strong structural variance in the pre-processed data. Note that the fit percentage f of a three-mode principal component solution can be computed as

$$f = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\hat{x}_{ijk}^2}{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}x_{ijk}^2}, \tag{3}$$

where

$$\hat{x}_{ijk} = \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R}a_{ip}b_{jq}c_{kr}g_{pqr} \tag{4}$$

(Kroonenberg, 1983, pp. 81–82). Subsequently, the residuals of these original analyses were further analysed by means of Tucker3. We opted for Tucker3 analyses, even when the original analysis was a Parafac analysis, because the Parafac restriction on the core array $\underline{\mathbf{G}}$ in (1) is a very strong constraint that will most probably not hold once the strong Parafac structure has been removed. In particular, using Tucker's (1966) quick procedure for computing approximate Tucker3 fit percentages (compare Kiers and der Kinderen, 2003), we obtained approximations f_{res} for the percentage of variance of the residuals that is accounted for by Tucker3 models of three different complexities (P,Q,R): (2,2,2), (2,3,4), and (4,4,4). In Figure 1, the resulting f_{res}-values are represented by stars. Moreover, to show that the residuals contain more Tucker3 structure than random noise, we sampled 1,000 data sets of the same size as the real data sets from the standard normal distribution and computed approximate Tucker3 fit percentages f_{noise} for all these random data sets in the above mentioned three complexities. For each real data set and each complexity, the observed f_{noise}-values for the 1,000 random data sets are represented by dots in Figure 1.

Table 1. Details of the original analysis of the six data sets: Size of the data set, type, complexity and fit percentage f_{strong} of the original analysis, and the reference for the original article.

		Original analysis			
Data set	Size	Type	Complexity	f_{strong}	Reference
Cars and stars	$25 \times 39 \times 34$	Parafac*	4	29.2	Harshman and De Sarbo (1984)
Chopin	$20 \times 24 \times 38$	Tucker3	(2,2,1)	40.8	Ceulemans and Kiers (2006)
Anxiety	$140 \times 14 \times 11$	Tucker3	(6,4,3)	41.1	Kiers and Van Mechelen (2001)
Bread	$10 \times 11 \times 8$	Parafac	2	49.2	Bro (1998)
Osgood	$15 \times 10 \times 6$	Tucker3	(3,3,2)	72.2	Kiers (1998)
Tongue	$10 \times 13 \times 5$	Parafac	2	92.6	Harshman, Ladefoged, and Goldstein (1977)

* One of the component matrices of the Parafac solution was constrained to be orthogonal.

2.2. Comparison of the approximate Tucker3 fit percentages for the residual data and for the randomly sampled data

Comparing the observed f_{noise}-values for the 1000 randomly sampled data sets with the f_{res}-value for the residual data clearly shows that f_{res} is always considerably larger than

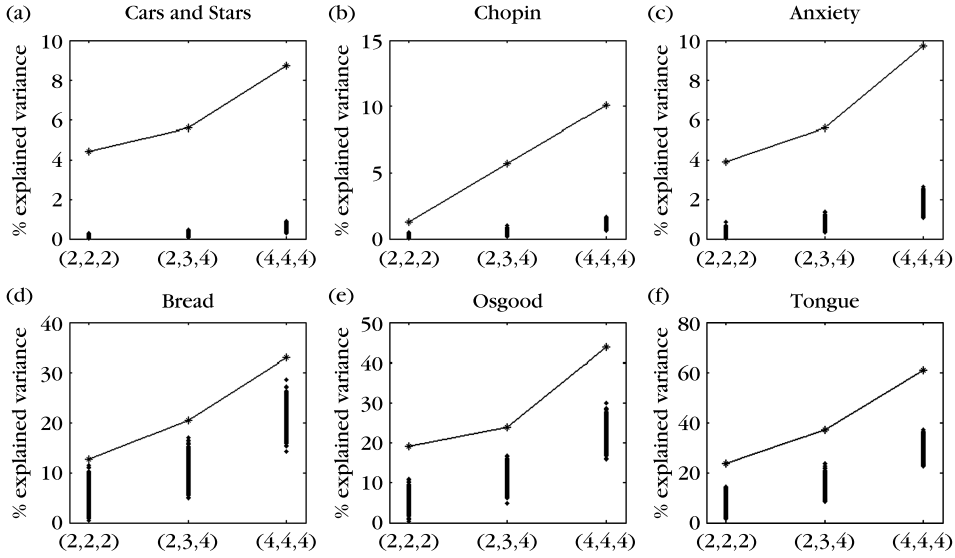

Figure 1. f_{noise}-values for 1,000 randomly sampled data sets and f_{res}-values for the six real data sets, for three different Tucker3 complexities: (2,2,2), (2,3,4), and (4,4,4). The f_{noise}-values are represented by dots, the f_{res}-values by stars which are connected by a line for each data set.

the corresponding f_{noise}-values. This result indicates that the residuals of real three-way three-mode data sets contain more Tucker3 structure than random noise does. As such, they corroborate our conjecture that simulation studies in which data sets are generated by adding randomly sampled noise to a three-mode principal component structure of low complexity are unrealistic.

As an aside, it is interesting to note that f_{noise} and f_{res} are related to the size of the data sets: The smaller the data set, the larger f_{noise} and f_{res}. This finding implies that the information in data sets with no strong underlying Tucker3 or Parafac structure is easier to summarize in case of small data sets than in case of larger data sets, as could be expected.

2.3. Division of structural variance over strong and weak three-mode principal component structures

To gain insight into how much variance of the pre-processed real data sets can generally be explained by weak three-mode principal component structures of different complexities, we calculated weak fit percentages f_{weak} as follows:

$$\begin{aligned} f_{weak} &= \frac{\text{expl.var.(weak components)}}{\text{total variance}} \times 100 \\ &= \frac{\text{expl.var.(weak components)} \times 100}{\text{total variance of residuals}} \times \frac{\text{total variance of residuals}}{\text{total variance}} \\ &= f_{res} \times \frac{100 - f_{strong}}{100}. \end{aligned} \tag{5}$$

Figure 2 represents, for each of the six real data sets and for each complexity of the Tucker3 analysis of the residuals, the amounts of total variance that are due to the strong

structure (f_{strong}) and the weak structure (f_{weak}) by black and white bars respectively. From this figure, it can be seen that these f_{weak}-values roughly vary between 1 and 17% and clearly increase for more complex Tucker3 analyses of the residuals. The f_{weak}-values are not so useful for designing the simulation studies of the present paper, however, as they do not shed light on how the structural variance of real data sets is divided between strong and weak three-mode principal component models. Therefore, we also computed proportions f_{weak}/f_{strong}, displayed in Table 2. Except for the bread data, these proportions vary between .02 and .11 for complexity (2,2,2), between .03 and .14 for complexity (2,3,4), and between .05 and .21 for complexity (4,4,4), with mean values of .06, .08, and .14 respectively. Note that the much larger values for the bread data are probably caused by the presence of a third strong Parafac-component in these data, as is also suggested by the CORCONDIA results reported by Bro and Kiers (2003). Indeed, reanalysing the bread data under the assumption of three strong Parafac components yielded proportions f_{weak}/f_{strong} of .06, .13 and .21 for the three complexities of the weak Tucker3 analyses, respectively, which all lie within the above-mentioned ranges.

3. Model selection heuristics

In this section, we describe the basic principle of the two model selection heuristics under evaluation: the numerical convex hull based method (Ceulemans & Van Mechelen, 2005; Ceulemans & Kiers, 2006) and CORCONDIA (Bro & Kiers, 2003).

3.1. Numerical convex hull based heuristic

The numerical convex hull based heuristic of Ceulemans and Van Mechelen (2005) and Ceulemans and Kiers (2006) builds on the visual inspection based model selection method proposed by Kroonenberg and Van der Voort (1987), Kroonenberg and Oort (2003) and Murakami and Kroonenberg (2003). The latter method boils down to visually inspecting scree-like plots with a measure of the fit (e.g. residual sum of squares, variance accounted for) of the solutions amongst which one wants to choose on the *Y*-axis and a measure of their complexity (e.g. degrees of freedom, number of free parameters, sum of the number of components for the three modes) on the *X*-axis. The key principle of the method is that one should favour the solutions that are on or

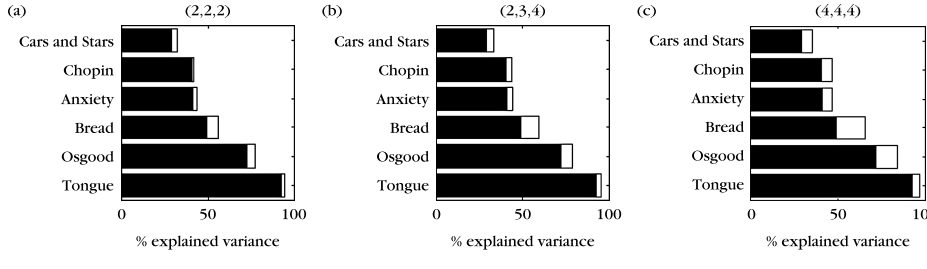

Figure 2. f_{strong}- and f_{weak}-values for each of the six real data sets and for each complexity of the Tucker3 analysis of the residuals. The f_{strong}- and f_{weak}-values are represented by the black and the white bars, respectively.

Table 2. f_{weak}/f_{strong}-values for the six real data sets, for three different Tucker3 complexities: (2,2,2), (2,3,4), and (4,4,4).

Data set	(2,2,2)	(2,3,4)	(4,4,4)
Cars and stars	.11	.14	.21
Chopin	.02	.08	.15
Anxiety	.06	.09	.14
Bread	.13	.21	.34
Osgood	.07	.09	.17
Tongue	.02	.03	.05

close to an elbow in the lower (in case the fit-measures quantify the badness of fit of the solutions) or higher (in case the fit-measures quantify the goodness of fit) boundary of the convex hull of this plot, because these solutions have the best balance of fit and complexity.

The contribution of Ceulemans and Van Mechelen (2005) and Ceulemans and Kiers (2006) consisted of providing a fully programmable version of the visual inspection based method by proposing a procedure which numerically assesses the higher or lower boundary of the convex hull of a fit versus complexity plot as well as the location of the elbow in this boundary. Assuming that the fit measures quantify the goodness of fit of the solutions, this procedure can be summarized in eight steps.

The first six steps serve to find the solutions on the higher boundary of the convex hull:

1. Determine the complexity value c and fit value f of all solutions from which one wishes to choose.
2. For each of the n observed complexity values c, retain only the best-fitting solution.
3. Sort the n retained solutions by their complexity values c from least to most complex and denote them by s_i $(i = 1, \ldots, n)$.
4. Exclude all solutions s_i for which a solution s_j $(j < i)$ exists such that $f_j > f_i$.
5. Consecutively consider all triplets of adjacent solutions. Exclude the middle solution if its point is located below or on the line connecting its neighbours in a fit value f versus complexity value c plot.
6. Repeat step 5 until no solution can be excluded. The remaining solutions all lie on the higher boundary of the convex hull.

The seventh step consists of determining the *st*-values of the 'hull' solutions obtained, with

$$st_i = \frac{f_i - f_{i-1}}{c_i - c_{i-1}} \Big/ \frac{f_{i+1} - f_i}{c_{i+1} - c_i}. \tag{6}$$

A relatively large *st*-value indicates that the solution of complexity c_i fits the data considerably better than the solution of complexity c_{i-1}, whereas the fit values of the solutions of complexity c_i and c_{i+1} differ much less. Finally, in the eighth step, the solution with the highest *st*-value is selected, which is the solution after which the increase in fit levels off.

For an extensive theoretical and empirical comparison of the numerical convex hull based heuristic and Timmerman and Kiers's (2000) DIFFIT method, we refer to Ceulemans and Kiers (2006). For this paper, it suffices to note that the 'hull' heuristic outperforms the DIFFIT method.

To evaluate the performance of the numerical convex hull based heuristic in discriminating between strong and weak three-mode principal component structures, we have to select measures to indicate the fit and complexity of the Tucker3 and Parafac solutions under consideration. As a fit measure we will use the fit percentage which can be calculated according to (3). The complexity of a Parafac solution will be quantified by the number of components R. Furthermore, the complexity of a Tucker3 solution will be expressed by computing the sum of the number of components P, Q, and R for the three modes. Note that the latter quantification of the complexity of Tucker3 solutions has been shown to yield slightly better model selection results than the use of the number of free parameters (Ceulemans & Kiers, 2006).

To illustrate the use of the numerical convex hull based heuristic, Figure 3 shows a goodness-of-fit versus complexity plot of 498 Tucker3 solutions for a simulated Tucker 3 data set (see Section 4) where both the strong and weak structures were of complexity (2,2,2). Applying the numerical convex hull based model selection procedure to this plot showed that five solutions lie on the higher boundary of the convex hull, which is represented by a line in Figure 3. The scree test values *st* of these five 'hull' solutions are given in Table 3. From this table, one may conclude that the 'hull' heuristic indicates the selection of the Tucker3 solution of complexity (4,4,4).

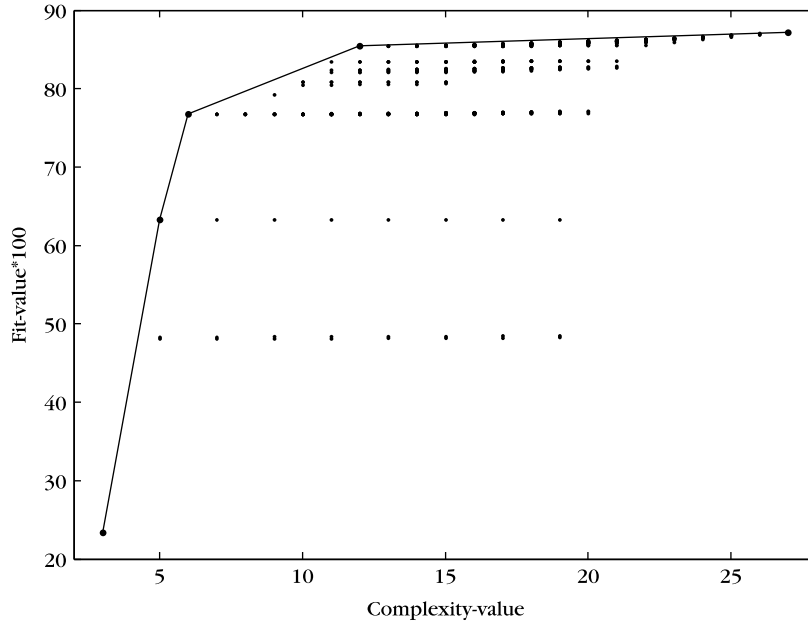

Figure 3. Plot of goodness-of-fit versus complexity for the 498 Tucker3 solutions for a simulated Tucker3 data set in which both the strong and weak structures were of complexity (2,2,2). The line represents the higher boundary of the convex hull and the larger points indicate the five 'hull' solutions.

Table 3. Goodness-of-fit-values *f*, complexity values *c* and scree test values *st* of the five Tucker3 solutions on the higher boundary of the convex hull of Figure 3.

(*P*,*Q*,*R*)	*c*	*f*	*st*
(1,1,1)	3	.2723	–
(1,2,2)	5	.7400	3.69
(2,2,2)	6	.8034	7.08
(4,4,4)	12	.8571	7.30
(9,9,9)	27	.8755	–

3.2. CORCONDIA

CORCONDIA, proposed by Bro and Kiers (2003), is a method for choosing among Parafac solutions of different complexities for a given data set. In particular, this method selects the most complex Parafac model that is still *appropriate*. Appropriate here means that the components primarily reflect trilinear variation in the data. This implies that omitting the Parafac superidentity constraint on the core array $\underline{\mathbf{G}}$ in (1), and hence allowing for interactions among the components from the different modes, would not 'significantly' increase the fit of the model to the data. As such, given specific component matrices **A**, **B** and **C** as obtained by applying the Parafac algorithm, the principle of CORCONDIA boils down to comparing the entries of an unconstrained Tucker3 core array with those of a superidentity array $\underline{\mathbf{T}}$ (i.e. $t_{pqr} = 1$ if $p = q = r$, and $t_{pqr} = 0$ otherwise):

$$\text{core consistency} = 100\left(1 - \frac{\sum_{r_1}^{R}\sum_{r_2}^{R}\sum_{r_3}^{R}\left(g_{r_1r_2r_3}^{T3} - t_{r_1r_2r_3}\right)^2}{R}\right), \tag{7}$$

where $\underline{\mathbf{G}}^{T3}$ is obtained as follows:

$$\text{vec}\mathbf{G}^{T3} = \left(\mathbf{C}\otimes\mathbf{B}\otimes\mathbf{A}\right)^{+}\text{vec}\mathbf{X}, \tag{8}$$

with the superscript '+' indicating a pseudoinverse. The statistic in (7) is called the core consistency as it quantifies the extent to which the Parafac assumptions hold for the Tucker3 core array. Specifically, a core consistency value close to 100% indicates that the Tucker3 core array almost takes the form of a superidentity array and, hence, that the Parafac solution under study is appropriate. A core consistency value of 50% or lower means that the off-diagonal entries of the Tucker3 core array are also rather large, implying that the Parafac solution is inappropriate. When computing the core consistency values of Parafac solutions of increasing complexity for a given data set, these values will typically decrease more or less monotonically and slowly with the number of components as long as the respective solutions are not too complex. However, once too many components are fitted, the core consistency will decrease more dramatically, because some components will describe non-trilinear variation, leading to high off-superdiagonal entries in the Tucker3 core array. As such, Bro and Kiers's (2003) recommendation to choose the most complex Parafac model that is still appropriate boils down to selecting the Parafac solution for which adding one more component drastically decreases the core consistency.

In the present paper, we wish to evaluate the performance of CORCONDIA in an extensive simulation study, without having to manually compare the core consistency values of different Parafac solutions for the same data set. Therefore, we propose to formalize the CORCONDIA heuristic as follows:

1. Determine the core consistency value *cc* of all Parafac solutions from which one wishes to choose.
2. For each solution s_R, where R indicates the number of components considered, determine the d_R-value which quantifies how drastically the core consistency value *cc* decreases when one more component is added:

$$d_R = \frac{cc_R - cc_{R+1}}{cc_{R-1} - cc_R}. \tag{9}$$

   Note that if $cc_{R-1} - cc_R \leq 0$, the nominator of (9) is set to a small value, say, 0.01. Furthermore, negative *cc*-values are set to 0.
3. Select the solution s_R with the highest d_R-value, for which $cc_R \geq 80$, indicating that the solution is sufficiently appropriate. Note that the cut-off value of 80 was chosen on the basis of the core consistency values that were obtained for real data sets by Bro and Kiers (2003). Pilot studies showed that setting the cut-off value to 90 or 70 has little effect on the results.

To illustrate the use of the CORCONDIA heuristic, Table 4 shows the core consistency values *cc* and the d_R-values of nine Parafac solutions for a simulated Parafac data set (see Section 4) where the strong Parafac structure and the weak Tucker3 structure were of complexity 2 and (2,3,4), respectively. From this table, one may conclude that adding a fifth Parafac component drastically decreases the core consistency; hence, the selection of the Parafac solution of complexity 4 is indicated.

Table 4. Core consistency values cc and d_R-values of nine Parafac solutions for a simulated Parafac data set where the strong Parafac structure and the weak Tucker3 structure were of complexity 2 and (2,3,4), respectively.

Complexity	*cc*	d_R
1	100	-
2	100	5.50
3	99.95	17.05
4	99.01	105.64
5	−166.94	-
6	−2037.54	-
7	−215.58	-
8	−1557.58	-
9	−2872.01	-

4. Simulation studies

In this section, we present two extensive simulation studies in which we evaluate the extent to which the numerical convex hull based model selection heuristic succeeds in disentangling strong and weak underlying three-mode principal component structures in Tucker3 and Parafac analysis. In particular, we investigate how often the hull heuristic

selects a solution of the strong complexity. Moreover, we examine whether and how four data characteristics influence its performance: size of the data array, complexity of the strong structure, complexity and importance of the weak structure, and amount of random noise on the data. Finally, in the Parafac simulation study, we also compare the performance of the hull and CORCONDIA heuristics.

4.1. Design of the Tucker3 and Parafac simulation studies

In the Tucker3 as well as in the Parafac simulation study, we constructed data arrays $\underline{\mathbf{X}}$ by systematically varying four independent variables:

1. The size $I \times J \times K$ of the data arrays at three levels: $200 \times 10 \times 10, 50 \times 20 \times 20$, $27 \times 27 \times 27$.
2. The complexity (P_s,Q_s,R_s) and R_s of the Tucker3 and Parafac models that constitute the strong structure of the data arrays, at three levels: for Tucker3, (2,2,2), (2,3,4), (4,4,4); and for Parafac, 2, 3, 4.
3. The complexity (P_w,Q_w,R_w) of the Tucker3 model that constitutes the weak structure of the data arrays and the percentage α of structural variance that it explains, at nine levels: (P_w,Q_w,R_w)=(2,2,2) and $\alpha = 2\%$; (2,2,2) and 6%, (2,2,2) and 10%; (2,3,4) and 6%; (2,3,4) and 10%; (2,3,4) and 14%; (4,4,4) and 10%; (4,4,4) and 14%; (4,4,4) and 18%. Note that the weak structure was always a Tucker3 structure, even when the strong structure is a Parafac structure, because the superidentity constraint on the Tucker3 core array $\underline{\mathbf{G}}$ in (1) is a very strong constraint that will probably not hold for the weak structure of most three-way three-mode data sets, unless this weak structure is of complexity (1,1,1) or (2,2,2). However, as the simulated weak structures of complexity (2,2,2) will often be fitted very well by a Parafac model with two components, the corresponding conditions of our simulation studies can also be considered as conditions in which the weak structure is a Parafac structure. Furthermore, note that the α levels were selected on the basis of the proportions $f_{\text{weak}}/f_{\text{strong}}$ that were found in our reanalyses of real data sets in Section 2. Indeed, as α indicates the percentage of weak structural variance, α levels of 2, 6, 10, 14, and 18% boil down to $(f_{\text{weak}})/(f_{\text{strong}}) = (\%\text{ weak structural variance})/(\%\text{ strong structural variance}) = \alpha/(100 - \alpha)$ proportions of .02, .06, .11, .16, and .22, respectively.
4. The amount β of error in the data, at three levels: 15, 30, 45%.

In both simulation studies, we considered 100 replicates for each cell of the design, resulting in 24,300 Tucker3 and 24,300 Parafac data arrays (3 sizes $\times$ 3 strong structures $\times$ 9 weak structures $\times$ 3 amounts of error $\times$ 100 replications).

4.1.1. Construction and analysis of Tucker3 data sets

The 24,300 Tucker3 data sets were constructed by

$$\mathbf{X} = \mathbf{A}_s\mathbf{G}_s(\mathbf{C}_s'\otimes\mathbf{B}_s') + \mathbf{A}_w\mathbf{G}_w(\mathbf{C}_w'\otimes\mathbf{B}_w') + \mathbf{E}. \tag{10}$$

In (10), $\mathbf{A}_s\mathbf{G}_s(\mathbf{C}_s'\otimes\mathbf{B}_s')$, $\mathbf{A}_w\mathbf{G}_w(\mathbf{C}_w'\otimes\mathbf{B}_w')$ and $\mathbf{E}$ constitute the strong Tucker3 structure, the weak Tucker3 structure, and random noise, respectively. The component matrices $\mathbf{A}_s$, $\mathbf{A}_w$, $\mathbf{B}_s$, $\mathbf{B}_w$, $\mathbf{C}_s$, and $\mathbf{C}_w$ are random orthonormal matrices, the core matrices $\mathbf{G}_s$ and

$\mathbf{G}_w$ are sampled from a uniform distribution with entries ranging from -0.5 to $+0.5$, and the error matrix $\mathbf{E}$ is sampled from the standard normal distribution. $\mathbf{E}$ is rescaled so as to represent $\beta\%$ of the variance of $\mathbf{X}$. Similarly, the weak structural part $\mathbf{A}_w\mathbf{G}_w(\mathbf{C}'_w \otimes \mathbf{B}'_w)$ is rescaled so as to account for $\alpha\%$ of the structural variance, that is, $\alpha\%$ of the variance of $\mathbf{X} - \mathbf{E}$. Note that $\mathbf{X}$ ($I \times JK$) and $\mathbf{G}_s$ and $\mathbf{G}_w$ (both $P \times QR$) are matricized versions of the corresponding three-way arrays.

This construction method may sometimes lead to data sets of which the strong or weak structural parts, $\mathbf{A}_s\mathbf{G}_s(\mathbf{C}'_s \otimes \mathbf{B}'_s)$ or $\mathbf{A}_w\mathbf{G}_w(\mathbf{C}'_w \otimes \mathbf{B}'_w)$, can be fitted almost as well by solutions with a lower sum of the number of components for the three modes than $P_s + Q_s + R_s$ and $P_w + Q_w + R_w$, respectively. The latter is undesirable for the present research, as it hampers a fair evaluation of the performance of the numerical convex hull heuristic. Indeed, if $\mathbf{A}_s\mathbf{G}_s(\mathbf{C}'_s \otimes \mathbf{B}'_s)$ can be fitted almost perfectly with less than $P_s +$ $_s + Q_s + R_s$ components, this means that the strong structural part is in practice partly weak. Furthermore, for those cases in which the hull heuristic fails to disentangle the strong and weak structures correctly, it is interesting to examine how often the method selects a solution that retrieves the strong as well as the weak underlying structure, that is, a solution of complexity ($P_s + P_w, Q_s + Q_w, R_s + R_w$). This can only be checked if all the components of the weak structure contribute 'significantly' to the weak structural variance. Therefore, while constructing the Tucker3 data arrays, all data arrays were discarded for which more than 95% of the strong or weak structural sum of squares can be fitted by Tucker3 solutions of lower complexity than the underlying strong and weak solutions.

For each of the resulting 24,300 simulated data sets, sums of the number of components for the three modes and approximate fit percentages (Kiers and der Kinderen, 2003) were obtained for 498 Tucker3 solutions of complexity (1,1,1) to (9,9,9). Subsequently, the numerical convex hull based model selection heuristic was applied to these 498 sums and fit percentages. With respect to the number of Tucker3 solutions considered, note that Wansbeek and Verhees (1989) proved that a Tucker3 model for which the number of components for, for instance, the first mode exceeds the product of the number of components for the other two modes can be reduced to a Tucker3 model for which the number of first mode components equals the product of the number of components for the other two modes; this implies that of all 729 Tucker3 solutions the 231 solutions for which $P > QR$, $Q > PR$, or $R > PQ$ can be omitted.

4.1.2. Construction and analysis of Parafac data sets

In the Parafac simulation study, the 24,300 data sets were constructed in the same way as in the Tucker3 simulation study; however, sampling $\mathbf{A}_s$, $\mathbf{B}_s$, and $\mathbf{C}_s$ from the standard normal distribution and fixing $\mathbf{G}_s$ such that rewriting $\mathbf{G}_s$ into $\underline{\mathbf{G}}_s$ yields a superidentity array. Each of the resulting simulated data sets was analysed with one rationally started run of the Parafac algorithm with the number of components R varying from 1 to 9. Each of these nine analyses per data set yielded a complexity value c equal to R, a fit percentage f, and a core consistency value cc as computed by (7). Finally, model selection was performed on the basis of the numerical convex hull based heuristic as well as using CORCONDIA.

4.2. Tucker3 results

Overall, the hull heuristic correctly discriminated between the strong and weak underlying Tucker3 structures in 63.5% of the 24,300 Tucker3 cases. In other words, for

15,426 data sets the hull heuristic selected a model of the same complexity as the strong underlying Tucker3 structure. To investigate how this overall performance of the hull heuristic is qualified by the four manipulated data characteristics, we conducted an analysis of variance with four between-block factors (size of the data array, complexity of the strong Tucker3 structure, complexity and importance of the weak Tucker3 structure and amount of error) and with a variable coding whether or not the hull heuristic succeeded in disentangling the strong and weak underlying Tucker3 structures as the dependent variable. Only considering effects that account for at least 5% of the variance of the dependent variable (i.e. effects for which the intra-class correlation $\hat{\rho}_I \geq .05$), the analysis revealed main effects of complexity of the strong Tucker3 structure ($\hat{\rho}_I = .10$), complexity and importance of the weak Tucker3 structure ($\hat{\rho}_I = .30$) and amount of random noise ($\hat{\rho}_I = .10$). The main effects for complexity of the strong Tucker3 structure and amount of error imply that the more complex the strong underlying Tucker3 solution and the less error in the data, the more difficult it is for the hull heuristic to discriminate between the strong and weak underlying Tucker3 structures (see Figures 4(a) and 4(b)). Regarding the main effect of the complexity and importance of the weak structure, it can be derived from Figure 4(c) that the performance is very similar for the three weak complexities, when averaging over the three corresponding importance levels (i.e. the three proportions of correct selection of the strong complexity amount to .66, .68 and .56). As such, the main difference among the three weak complexities pertains to the extent to which the performance of the hull heuristic is qualified by the corresponding importance levels: The lower the weak complexity, the more the performance of the hull heuristic deteriorates with increasing importance of the weak underlying structure.

To gain further insight in which solutions are selected for the 8,874 Tucker3 data sets for which the hull heuristic failed to retrieve the complexity of the strong underlying Tucker3 structure, we examined for how many of those cases this heuristic selected a solution that incorporates both the strong and weak underlying structures, that is, a solution of complexity ($P_s + P_u$, $Q_s + Q_u$, $R_s + R_w$). This was the case for 7,349 of those data sets. As such, it can be concluded that in more than 93% of the 24,300 data sets, the hull heuristic either extracts the strong underlying Tucker3 structure from the data or the full combination of the strong and weak underlying Tucker3 structures.

Finally, it should be noted that for 4,539 (51.2%) of the 8,874 'incorrect selection' cases, the solution of complexity (P_s, Q_s, R_s) had the second highest *st*-value. For an example, see Figure 3 and Table 3, the contents of which were described in Section 3.1. These results qualify the term 'incorrect selection', because they imply that if the numerical convex hull based procedure had been applied in a somewhat relaxed way (i.e. also considering solutions with a high but not maximal scree test value *st*), as one will often do in practice (see Ceulemans and Kiers, 2006), the solution of complexity (P_s, Q_s, R_s) would in at least 50% of the cases also have been selected for further consideration.

4.3. Parafac results

4.3.1. Numerical convex hull based heuristic

Overall, the hull heuristic succeeds in disentangling the strong underlying Parafac structure from the weak Tucker3 structure in 16,607 of the 24,300 simulated data sets, in that it selects a Parafac model which is equally complex as the strong Parafac

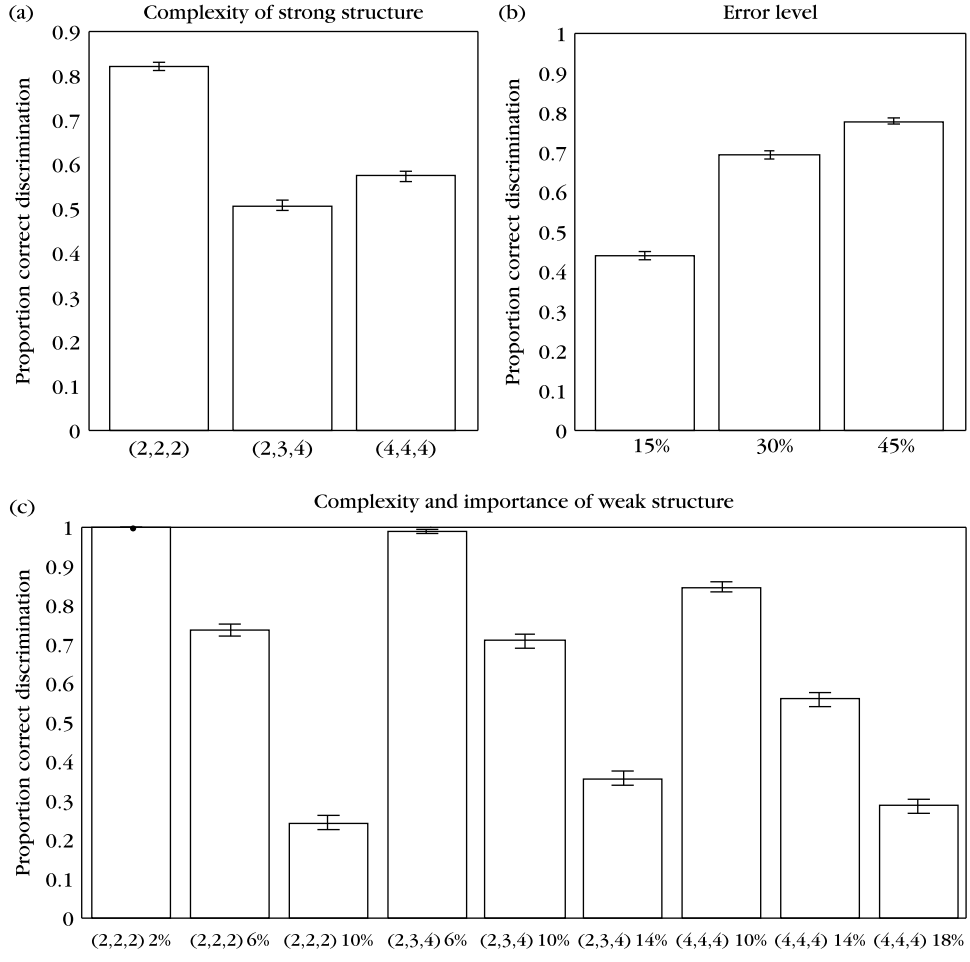

Figure 4. Proportion of correct discrimination between the strong and weak underlying Tucker3 structures for the 24,300 Tucker3 data arrays and associated 95% confidence intervals, as a function of (a) complexity of the strong Tucker3 structure, (b) error level and (c) complexity and importance of the weak Tucker3 structure.

structure. Again, an analysis of variance was performed with the four manipulated data characteristics as independent variables and with a dependent variable coding whether or not the hull heuristic selected a solution of the same complexity as the strong underlying Parafac structure. This analysis of variance yielded a main effect of the complexity of the strong Parafac structure ($\hat{\rho}_I = .09$): the more complex the strong Parafac structure, the harder it becomes for the hull heuristic to disentangle it from the weak structure (see Figure 5(a)). Furthermore, an important main effect of the complexity and importance of the weak Tucker3 structure ($\hat{\rho}_I = .30$) was revealed: for the three weak complexities considered, the hull heuristic performs about equally well when considering the lowest importance level only. However, taking all importance levels into account, the lower the weak complexity, the more the performance of the hull heuristic deteriorates with increasing importance of the weak underlying structure (see Figure 5(b)).

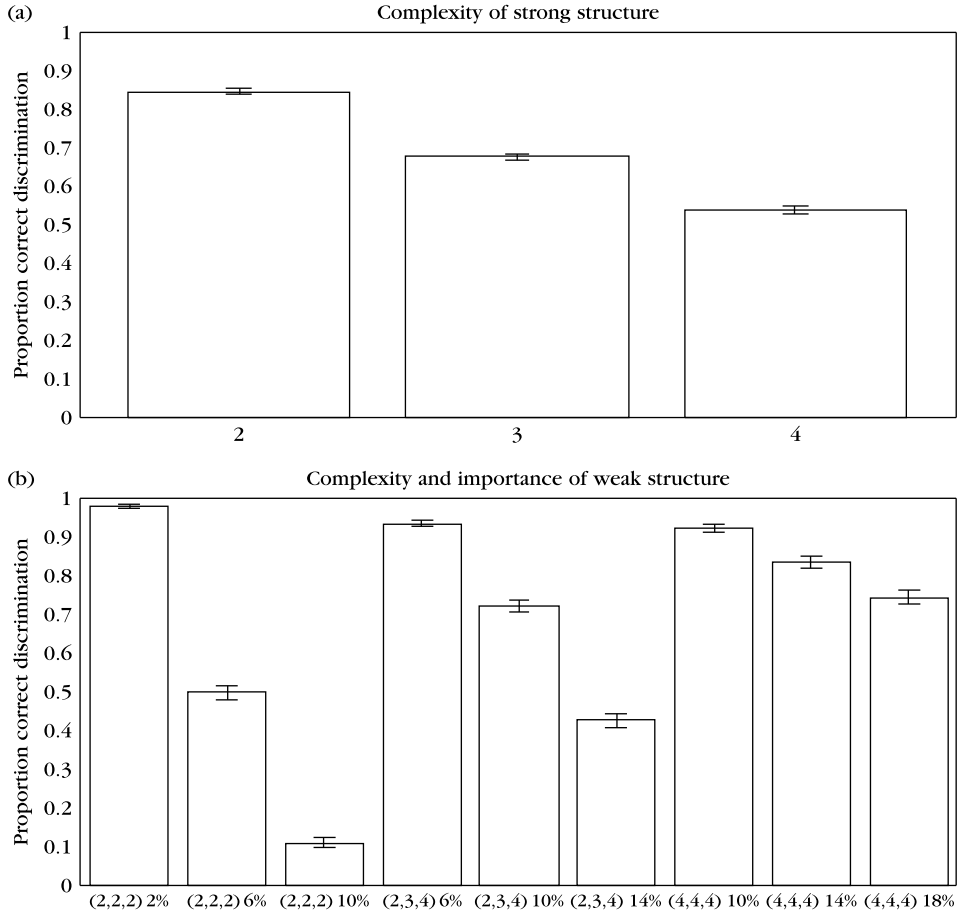

Figure 5. Proportion of correct discrimination between the strong and weak three-mode principal component structures underlying the 24,300 Parafac data arrays and associated 95% confidence intervals, as a function of (a) complexity of the strong Parafac structure and (b) complexity and importance of the weak Tucker3 structure.

We further studied what kind of solution was selected for the 7,693 data sets in which the hull heuristic fails to differentiate between the strong and weak underlying three-mode principal component structures. Specifically, we examined whether the hull heuristic tends to select a solution with more than R_s components in such cases, implying that some of the weak structural variation is modelled, or whether it will instead choose a solution with less than R_s components. In this respect, it was found that in the vast majority of the cases (i.e. 6,209) the hull heuristic indicates the selection of a solution with more components than the strong underlying Parafac solution. For instance, for 3,673 of the 3,831 data sets where the complexity of the weak structure equalled (2,2,2), a Parafac solution was selected with $R_s + 2$ components. As such, it can be concluded that in more than 93% of the 24,300 Parafac data sets, the hull heuristic either retreives the strong underlying Parafac structure from the data or a combination of this strong structure and a part of the weak Tucker3 structure.

As in the Tucker3 simulation study, we checked for how many of the 7,693 'incorrect selection' cases the solution of complexity R_s had the second highest *st*-value, implying

that in practical applications of the hull procedure this solution would also have been retained for further consideration. This was the case for 5,145 (67%) of these 7,693 data sets.

4.3.2. CORCONDIA

CORCONDIA performed clearly worse than the hull heuristic. Specifically, CORCONDIA succeeded for only 3,741 of the 24,300 simulated Parafac data sets (15.4%) in discriminating the strong underlying Parafac structure from the weak Tucker3 structure. An analysis of variance with the four manipulated data characteristics as independent variables and with the performance of CORCONDIA as the dependent variable, revealed only a main effect ($\hat{\rho}_I = .08$) of the size of the data sets: For sizes $200 \times 10 \times 10$, $50 \times 20 \times 20$ and $27 \times 27 \times 27$ CORCONDIA indicated the selection of a solution with R_s components in 2250, 821 and 670 of the cases, respectively; thus, it can be concluded that the more asymmetric the data array, the better the performance of CORCONDIA.

Investigating the results in more detail shows that CORCONDIA systematically (i.e. in 18,938 or 77.9% of the cases) selects a solution with more components than the strong underlying Parafac solution, where the solutions with $R_s + 1$ (10,632 cases or 43.8%) and $R_s + 2$ (6,512 cases or 26.8%) components are most often picked up. This implies that, even if the number of extracted Parafac components is larger than R_s, the core consistency values of the Parafac solutions are often larger than 80, indicating that the solutions are sufficiently appropriate.

5. Discussion

In this paper, it was argued that the simulation studies performed to evaluate the DIFFIT and numerical convex hull based model selection techniques (Ceulemans & Kiers, 2006; Kiers & der Kinderen, 2003; Timmerman & Kiers, 2000) were unrealistic in that these studies imply the assumption that the variance in any three-way three-mode data set is due to two sources only: (i) a (strong) three-mode principal component structure of low complexity; and (ii) random noise. Specifically, in line with earlier claims by Tucker *et al.* (1969) for factor analysis, it was argued that it is much more reasonable to assume the presence of an additional source of variance: a second, weak three-mode principal component structure that describes less prominent covariation in the data and is thus of less interest to the researcher. This claim was empirically validated by further analysing the residuals of six real and well-documented three-way data sets, demonstrating that these residuals contain more Tucker3 structure than randomly sampled data sets of the same size.

It is of interest to note that, for two-way data, it can also be expected that data always consist of a weak component structure, in addition to a strong one and purely random noise. This implies that methods for selection of numbers of components that are based on significance testing will, provided that n is sufficiently large, always lead to solutions including the weak components, because these are real (albeit unimportant) components. This demonstrates that selection of components should not simply be based on significance testing, but should take the actual importance of additional components, that is, the amount of structural variance explained, into account. Note that the amount of structural variance explained is a better criterion than the amount of total variance explained, because the amount of structural variance may be considerably lower than the amount of total variance due to the presence of a lot of noise.

Two new simulation studies were run according to the more realistic assumption that three-way three-mode data are constituted by the combination of a strong and a weak three-mode principal component structure plus random noise. The exact settings of these simulation studies were based on the reanalysis of the six real data sets. The results show that in about two-thirds of the simulated cases, the numerical convex hull based heuristic yields a model of the same complexity as the strong underlying structure, and, thus succeeds in disentangling strong and weak underlying structures. In the vast majority of the remaining third, this heuristic selects a solution that combines the strong structure and (part of) the weak structure; furthermore, the solution for the strong complexity has in about 51% and 67% of these Tucker3 and Parafac cases the second highest *st*-value, implying that in practice it will often also be retained for further consideration. Importantly, these findings indicate that the application of the hull procedure will almost never result in the disregarding of salient aspects of the data; additionally, the hull procedure will in some cases draw attention to less prominent but systematic relations in the data.

These overall results were further qualified by the manipulated data characteristics, implying that the performance of the hull heuristic in discriminating between strong and weak underlying structures deteriorates mostly with increasing complexity of the strong structure and increasing importance of the weak structure. This makes sense in that the increases in fit related to adding a component of the strong structure and a component of the weak structure become more similar when the strong structure is more complex or when the weak structure accounts for a larger part of the structural variance.

Additionally, the present research included a systematic comparison of the performance of the numerical convex hull based heuristic and CORCONDIA, from which it was concluded that CORCONDIA performs clearly worse than the hull heuristic in discriminating the strong Parafac structure in a data set from a weak Tucker3 structure. In particular, it was found that CORCONDIA is much more prone than the hull heuristic to extract too many components from the simulated data sets.

A limitation of the simulation studies presented is that the component matrices were sampled from normal distributions. Hence, a possible direction for future research would involve designing simulation studies in which the component matrices are sampled from other types of distributions. However, it should be noted that in the Ceulemans and Kiers (2006) simulation study the hull heuristic also performed well when the simulated data sets were based on empirically obtained component matrices.

Finally, as stated before, when analysing three-way data with Tucker3 and Parafac models, most researchers will want to focus on the strong structure. However, in some cases, it may be interesting to further analyse the residuals of the strong analysis (as was illustrated in Section 2) to gain insight into the weak structure that remains once the strong structure has been extracted. An important advantage of this sequential approach over a simultaneous modelling of the strong and weak structure is that the interactions between components of the strong structure and components of the weak structure are discarded, simplifying the interpretation of the solution obtained.

Acknowledgements

We thank Richard A. Harshman for giving us permission to use the cars and stars data as well as the tongue data; and Iven Van Mechelen for providing the anxiety data. The Chopin and Osgood data

can be downloaded from http://three-mode.leidenuniv.nl, the bread data from http://www.models.kvl.dk/research/data/. The authors are obliged to Jan Schepers for comments on an earlier version of this paper. The research reported in this paper was partially supported by the Research Council of K.U.Leuven (GOA/05/04).

References

Bro, R. (1998). *Multi-way analysis in the food industry. Models, algorithms, and applications*. PhD thesis, University of Amsterdam.

Bro, R., & Kiers, H. A. L. (2003). A new efficient method for determining the number of components in PARAFAC models. *Journal of Chemometrics*, *17*, 274–286.

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an *n*-way generalization of 'Eckart–Young' decomposition. *Psychometrika*, *35*, 283–319.

Ceulemans, E., & Kiers, H. A. L. (2006). Selecting among three-mode principal component models of different types and complexities: A numerical convex hull based method. *British Journal of Mathematical and Statistical Psychology*, *59*, 133–150.

Ceulemans, E., & Van Mechelen, I. (2005). Hierarchical classes models for three-way three-mode binary data: Interrelations and model selection. *Psychometrika*, *70*, 461–480.

Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika*, *19*, 149–161.

Harshman, R. A. (1970). Foundations of the PARAFAC procedure: Models and conditions for an explanatory multi-modal factor analysis. *UCLA Working Papers in Phonetics*, *16*, 1–84.

Harshman, R. A., & De Sarbo, W. S. (1984). An application of PARAFAC to a small sample problem, demonstrating pre-processing, orthogonality constraints and split-half diagnostic techniques. In H. G. Law, C. W. Snyder Jr, J. A. Hattie, & R. P. McDonald (Eds.), *Research methods for multimode data analysis* (pp. 602–642). New York: Praeger.

Harshman, R. A., Ladefoged, P., & Goldstein, L. (1977). Factor analysis of tongue shapes. *Journal of the Acoustical Society of America*, *62*, 697–707.

Kiers, H. A. L. (1998). Joint orthomax rotation of the core and component matrices resulting from three-mode principal components analysis. *Journal of Classification*, *15*, 245–263.

Kiers, H. A. L., & der Kinderen, A. (2003). A fast method for choosing the numbers of components in Tucker3 analysis. *British Journal of Mathematical and Statistical Psychology*, *56*, 119–125.

Kiers, H. A. L., & Van Mechelen, I. (2001). Three-way component analysis: Principles and illustrative application. *Psychological Methods*, *6*, 84–110.

Kroonenberg, P. M. (1983). *Three-mode principal component analysis: Theory and applications*. Leiden: DSWO.

Kroonenberg, P. M. (2008). *Applied multiway data analysis*. Hoboken, NJ: Wiley.

Kroonenberg, P. M., & De Leeuw, J. (1980). Principal component analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika*, *45*, 69–97.

Kroonenberg, P. M., & Oort, F. J. (2003). Three-mode analysis of multimode covariance matrices. *British Journal of Mathematical and Statistical Psychology*, *56*, 305–336.

Kroonenberg, P. M., & Van der Voort, T. H. A. (1987). Multiplicatieve decompositie van interacties bij oordelen over de werkelijkheidswaarde van televisiefilms [Multiplicative decomposition of interactions for judgements of realism of television films]. *Kwantitatieve Methoden*, *8*, 117–144.

Murakami, T., & Kroonenberg, P. M. (2003). Three-mode models and individual differences in semantic differential data. *Multivariate Behavioural Research*, *38*, 247–283.

Timmerman, M. E., & Kiers, H. A. L. (2000). Three-mode principal components analysis: Choosing the numbers of components and sensitivity to local optima. *British Journal of Mathematical and Statistical Psychology*, *53*, 1–16.

Tucker, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika*, *31*, 279–311.

Tucker, L. R., Koopman, R. F., & Linn, R. L. (1969). Evaluation of factor analytic research procedures by means of simulated correlation matrices. *Psychometrika*, *34*, 421–459.

Wansbeek, T., & Verhees, J. (1989). Models for multidimensional matrices in econometrics and psychometrics. In R. Coppi & S. Bolasco (Eds.), *Multiway data analysis* (pp. 543–552). Amsterdam: North Holland.

Received 4 June 2007; revised version received 1 September 2008