Low-Rank Approximation of Generic p x q x 2 Arrays and Diverging Components in the Candecomp/Parafac Model

Stegeman, Alwin

Published in:
SIAM Journal on Matrix Analysis and Applications

DOI:
10.1137/050644677

SIAM J. MATRIX ANAL. APPL.
Vol. 30, No. 3, pp. 988–1007

LOW-RANK APPROXIMATION OF GENERIC $p \times q \times 2$ ARRAYS AND DIVERGING COMPONENTS IN THE CANDECOMP/PARAFAC MODEL*

ALWIN STEGEMAN†

Abstract. We consider the low-rank approximation over the real field of generic $p \times q \times 2$ arrays. For all possible combinations of p, q, and R, we present conjectures on the existence of a best rank-R approximation. Our conjectures are motivated by a detailed analysis of the boundary of the set of arrays with at most rank R. We link these results to the Candecomp/Parafac (CP) model for three-way component analysis. Essentially, CP tries to find a best rank-R approximation to a given three-way array. In the case of $p \times q \times 2$ arrays, we show (under some regularity condition) that if a best rank-R approximation does not exist, then any sequence of CP updates will exhibit diverging CP components, which implies that several components are highly correlated in all three modes and their component weights become arbitrarily large. This extends Stegeman [*Psychometrika*, 71 (2006), pp. 483–501], who considers $p \times p \times 2$ arrays of rank $p+1$ or higher. We illustrate our results by means of simulations.

Key words. low-rank tensor approximations, border rank, arrays, Candecomp, Parafac, three-way arrays, degenerate Parafac solutions

AMS subject classifications. 15A03, 15A22, 15A69, 49M27, 62H25

DOI. 10.1137/050644677

1. Introduction.

We consider the problem of finding a best low-rank approximation to a three-way array $\underline{\mathbf{X}} \in \mathbb{R}^{p \times q \times 2}$. In this introductory section, we discuss the general problem of finding a best low-rank approximation to a k-way array and its applications in algebraic complexity theory (the multiplicative complexity of the computation of bilinear forms) and psychometrics (the Candecomp/Parafac (CP) model for three-way component analysis). Also, the consequences of an array not having a best low-rank approximation are discussed for these fields of research. Finally, we consider some results of the theory of matrix pencils with implications on the rank of $p \times q \times 2$ arrays, and show how our analysis fits into this literature.

1.1. Low-rank approximation of arrays.

Let the rank over a field $\mathcal{F}$ of a k-way array $\underline{\mathbf{X}} \in \mathcal{F}^{d_1 \times \cdots \times d_k}$ be defined in the usual way, i.e., as the smallest number of rank-1 arrays in $\mathcal{F}^{d_1 \times \cdots \times d_k}$ whose sum equals $\underline{\mathbf{X}}$; see Hitchcock [15, 16]. A k-way array has rank 1 over $\mathcal{F}$ if it is the outer product of k vectors in $\mathcal{F}^{d_1}, \ldots, \mathcal{F}^{d_k}$. The problem of finding a best rank-R approximation of $\underline{\mathbf{X}} \in \mathbb{R}^{d_1 \times \cdots \times d_k}$ boils down to minimizing

$$\left\| \underline{\mathbf{X}} - \sum_{r=1}^{R} \mathbf{a}_r^{(1)} \circ \cdots \circ \mathbf{a}_r^{(k)} \right\| \tag{1.1}$$

over the vectors $\mathbf{a}_r^{(j)} \in \mathcal{F}^{d_j}$, $j = 1, \ldots, k$, $r = 1, \ldots, R$, where $\circ$ denotes the outer product and $\|\cdot\|$ denotes some norm on $\mathcal{F}^{d_1 \times \cdots \times d_k}$. Unless stated otherwise, we will

*Received by the editors November 9, 2005; accepted for publication (in revised form) by L. De Lathauwer March 6, 2007; published electronically September 25, 2008. This research was supported by the Dutch Organisation for Scientific Research (NWO), VENI grant 451-04-102.
http://www.siam.org/journals/simax/30-3/64467.html
†Heijmans Institute of Psychological Research, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands (a.w.stegeman@rug.nl).

assume that $\mathcal{F} = \mathbb{R}$ and $\|\cdot\|$ is the Frobenius norm. We denote the rank of an array $\underline{\mathbf{X}}$ as $\text{rank}_\circ(\underline{\mathbf{X}})$.

For $k = 2$, all best rank-R approximations can be obtained from the singular value decomposition of the matrix to be approximated; see Eckart and Young [11]. However, for $k \geq 3$ a best rank-R approximation does not always exist. Examples of arrays that can be approximated arbitrarily well by arrays of lower rank are known from the algebraic complexity literature; see Bini et al. [1], Bini, Lotti, and Romani [2], and Bini [3], as well as from the psychometric and chemometric literature; see ten Berge, Kiers, and De Leeuw [37] and Paatero [30]. Stegeman [34] has shown, under some regularity condition, that generic $p \times p \times 2$ arrays of rank $p+1$ (a set of positive volume in $\mathbb{R}^{p\times p\times 2}$) do not have a best rank-p approximation. De Silva and Lim [10] show that a best rank-1 approximation always exists, while for any $k \geq 3$, any $d_1, \ldots, d_k \geq 2$, and any $R \in \{2, \ldots, \min(d_1, \ldots, d_k)\}$, a rank-$(R+1)$ array exists which has no best rank-R approximation. Also, [10] show that all $2 \times 2 \times 2$ arrays of rank 3 (a set of positive volume in $\mathbb{R}^{2\times 2\times 2}$) have no best rank-2 approximation, and that, for any $d_1, d_2, d_3 \geq 2$, the set of arrays in $\mathbb{R}^{d_1\times d_2\times d_3}$ which have no best rank-2 approximation has positive volume.

1.2. Algebraic complexity theory and array rank. An important problem in algebraic complexity theory is the multiplicative complexity of the computation of a set of bilinear forms $\mathbf{u}^T\mathbf{X}_k\mathbf{v}$, $k = 1, \ldots K$, where $\mathbf{u}$ and $\mathbf{v}$ are indeterminates and the $\mathbf{X}_k$ have elements in a field $\mathcal{F}$. Strassen [36] showed that the K bilinear forms $\mathbf{u}^T\mathbf{X}_k\mathbf{v}$ can be computed with R nonscalar multiplications (i.e., multiplications of two elements not in $\mathcal{F}$), where R is the rank over $\mathcal{F}$ of the array $\underline{\mathbf{X}}$ with slices $\mathbf{X}_k$. Indeed, if $\underline{\mathbf{X}} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r$, for vectors $\mathbf{a}_r$, $\mathbf{b}_r$, and $\mathbf{c}_r$ with elements in $\mathcal{F}$, then we have

$$\mathbf{u}^T\mathbf{X}_k\mathbf{v} = \sum_{r=1}^{R} (\mathbf{u}^T\mathbf{a}_r)\,(\mathbf{b}_r^T\mathbf{v})\,c_{kr}\,, \tag{1.2}$$

and the K bilinear forms $\mathbf{u}^T\mathbf{X}_k\mathbf{v}$ can be computed using R nonscalar multiplications. See also Brockett and Dobkin [6].

Suppose $\mathcal{F} = \mathbb{R}$ and $\underline{\mathbf{X}}$ can be approximated arbitrarily well by rank-$(R-1)$ arrays. In that case, we could replace $\underline{\mathbf{X}}$ in (1.2) by a rank-$(R-1)$ array close to it and use only $R-1$ nonscalar multiplications in the computation of the bilinear form. Since the rank-$(R-1)$ array can be chosen arbitrarily close to $\underline{\mathbf{X}}$, the error in the computation of the bilinear form can be made arbitrarily small. This idea was pointed out by Bini et al. [1], Bini, Lotti, and Romani [2], and Bini [3, 4] and has been used in the design of algorithms for matrix multiplication; see Bürgisser, Clausen, and Shokrollahi [7, Chapter 15] and the references therein. Hence, the nonexistence of a best rank-$(R-1)$ approximation of $\underline{\mathbf{X}}$ yields a faster and arbitrarily accurate computation of the bilinear forms.

To express the optimal computational gain that can be achieved by approximating $\underline{\mathbf{X}}$ with arrays of lower rank, Bini, Lotti, and Romani [2] have introduced the notion of *border rank*. The border rank of an array $\underline{\mathbf{X}}$, which we denote by $\text{rank}_B(\underline{\mathbf{X}})$, is defined as

$$\text{rank}_B(\underline{\mathbf{X}}) = \min\{R : \ \underline{\mathbf{X}} \text{ can be approximated arbitrarily well by arrays of rank } R\}\,. \tag{1.3}$$

From this definition it follows that $\text{rank}_B(\underline{\mathbf{X}}) \leq \text{rank}_\circ(\underline{\mathbf{X}})$. Results on the border rank of various arrays have been obtained by Bini [4, 5] and Landsberg [25]. For later

use we state the following result. Let $\text{rank}_i(\underline{\mathbf{X}})$ denote the rank of the set of mode i fibers of $\underline{\mathbf{X}}$, where a mode i fiber is a vector obtained by varying the mode i index and keeping all other indices fixed. This notion of rank is due to Hitchcock [15, 16] and the set of $\text{rank}_i(\underline{\mathbf{X}})$ for all i is called the *multilinear rank* of $\underline{\mathbf{X}}$ in De Silva and Lim [10]

PROPOSITION 1.1. *For a k-way array $\underline{\mathbf{X}} \in \mathbb{R}^{d_1 \times \cdots \times d_k}$, there holds*

$$\max_i \text{rank}_i(\underline{\mathbf{X}}) \leq \text{rank}_B(\underline{\mathbf{X}}) \leq \text{rank}_\circ(\underline{\mathbf{X}}) \,. \tag{1.4}$$

Proof. The second inequality follows from (1.3). A proof of the inequality $\text{rank}_i(\underline{\mathbf{X}}) \leq \text{rank}_\circ(\underline{\mathbf{X}})$ is given by De Silva and Lim [10]. We state it here for completeness. Let $R = \text{rank}_\circ(\underline{\mathbf{X}})$ and $\underline{\mathbf{X}} = \sum_{r=1}^{R} \mathbf{a}_r^{(1)} \circ \cdots \circ \mathbf{a}_r^{(k)}$. Then all mode i fibers lie in the span of $\mathbf{a}_1^{(i)}, \ldots, \mathbf{a}_R^{(i)}$. This implies $\text{rank}_i(\underline{\mathbf{X}}) \leq R$.

We show the first inequality by contradiction. Suppose $t = \text{rank}_B(\underline{\mathbf{X}}) < \text{rank}_i(\underline{\mathbf{X}})$. Then there exists a sequence of rank-t arrays $\underline{\mathbf{Y}}^{(n)}$ converging to $\underline{\mathbf{X}}$. But then also the matrices $\mathbf{Y}_i^{(n)}$, containing as columns the mode i fibers of $\underline{\mathbf{Y}}^{(n)}$, must converge to the matrix $\mathbf{X}_i$ containing the mode i fibers of $\underline{\mathbf{X}}$. This is a contradiction, since $\text{rank}(\mathbf{Y}_i^{(n)}) \leq t < \text{rank}(\mathbf{X}_i)$ for all n, and a matrix cannot be approximated arbitrarily well by matrices of lower rank. Note that the upper semicontinuity of the multilinear rank is used here; see De Silva and Lim [10]. ∎

1.3. The CP model and diverging components. Carroll and Chang [9] and Harshman [13] have independently proposed the same method for component analysis of three-way data arrays and named it Candecomp and Parafac, respectively. We denote the CP model as

$$\underline{\mathbf{X}} = \sum_{r=1}^{R} \omega_r \, (\mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r) + \underline{\mathbf{E}} \,, \tag{1.5}$$

where $\underline{\mathbf{X}}$ is a $d_1 \times d_2 \times d_3$ data array, ω_r is the weight of component r, and $\|\mathbf{a}_r\| = \|\mathbf{b}_r\| = \|\mathbf{c}_r\| = 1$ for $r = 1, \ldots, R$. The Frobenius norm of $\underline{\mathbf{E}}$ is minimized to find the R components $\mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r$ and the weights ω_r. For an overview and comparison of CP algorithms, see Hopke et al. [17] and Tomasi and Bro [41]. From (1.5) it is clear that the CP model tries to find a best rank-R approximation to the three-way array $\underline{\mathbf{X}}$.

The CP model (1.5) can be seen as a three-way extension of the principal component analysis model for matrices. For example, if the vectors $\mathbf{a}_r$ are interpreted as the components in mode 1, then $\mathbf{b}_r$ and $\mathbf{c}_r$ are the loadings on these components for modes 2 and 3, respectively. The real-valued CP model, i.e., where $\underline{\mathbf{X}}$ and the model parameters are real-valued, is used in a majority of applications in psychometrics and chemometrics; see Kroonenberg [21] and Smilde, Bro, and Geladi [33]. Complex-valued applications of CP occur in, e.g., signal processing and telecommunications research; see Sidiropoulos [32]. In this paper, we consider only the real-valued CP model.

A matrix notation of the CP model (1.5) is as follows. Let $\mathbf{X}_k$ $(d_1 \times d_2)$ and $\mathbf{E}_k$ $(d_1 \times d_2)$ denote the kth slices of $\underline{\mathbf{X}}$ and $\underline{\mathbf{E}}$, respectively. Then (1.5) can be written as

$$\mathbf{X}_k = \mathbf{A}\,\mathbf{C}_k\,\mathbf{\Omega}\,\mathbf{B}^T + \mathbf{E}_k \,, \qquad k = 1, \ldots K \,, \tag{1.6}$$

where $\mathbf{A}$ $(d_1 \times R)$ and $\mathbf{B}$ $(d_2 \times R)$ have the vectors $\mathbf{a}_r$ and $\mathbf{b}_r$ as columns, respectively, $\mathbf{\Omega}$ $(R \times R)$ is the diagonal matrix with the weights ω_r on its diagonal, and $\mathbf{C}_k$ $(R \times R)$

is the diagonal matrix with the kth elements of the vectors $\mathbf{c}_r$ on its diagonal. The model part of the CP model is characterized by $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{\Omega})$, where $\mathbf{C}$ $(d_3 \times R)$ has the vectors $\mathbf{c}_r$ as columns. We refer to $\mathbf{A}, \mathbf{B}, \mathbf{C}$ as the component matrices and to $\mathbf{\Omega}$ as the weights matrix.

The most attractive feature of CP is its uniqueness property. Kruskal [22] has shown that, for fixed residuals $\underline{\mathbf{E}}$, the vectors $\mathbf{a}_r$, $\mathbf{b}_r$, and $\mathbf{c}_r$ and the weights ω_r are unique up to sign changes and a reordering of the summands in (1.5) if

$$k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2\,R + 2\,, \tag{1.7}$$

where $k_{\mathbf{A}}$, $k_{\mathbf{B}}$, $k_{\mathbf{C}}$ denote the k-ranks of the component matrices. The k-rank of a matrix is the largest number x such that every subset of x columns of the matrix is linearly independent. Hence, contrary to the matrix principal components model, the CP components are rotationally unique if (1.7) holds.

However, the practical use of CP has been hampered by the occurrence of diverging CP components. In the majority of such cases, exactly two components displayed the following pattern. Let the model parameters of the nth update of a CP algorithm be denoted by a superscript (n). In the case of two diverging CP components, say s and t, the weights $\omega_s^{(n)}$ and $\omega_t^{(n)}$ become arbitrarily large in magnitude while the vectors $\mathbf{a}_s^{(n)}$, $\mathbf{b}_s^{(n)}$, and $\mathbf{c}_s^{(n)}$ become nearly identical (up to sign changes) to $\mathbf{a}_t^{(n)}$, $\mathbf{b}_t^{(n)}$, and $\mathbf{c}_t^{(n)}$ such that

$$\sum_{r=s,t} \omega_r^{(n)} \, (\mathbf{a}_r^{(n)} \circ \mathbf{b}_r^{(n)} \circ \mathbf{c}_r^{(n)})$$

remains "small." Hence, the contributions of components s and t diverge in nearly opposite directions, but their sum still contributes to a better fit of the CP model. The CP algorithm becomes very slow when this occurs; see Mitchell and Burdick [28]. When the CP algorithm is terminated, the CP components obtained are said to form a *degenerate CP solution*. Since this use of the term *degenerate* is different from its general meaning in mathematics, we will speak of *diverging CP components* instead. This also reflects the fact that this phenomenon occurs when running a CP algorithm, while a degenerate CP solution suggests a property of one CP solution only.

The first case of two diverging CP components was reported in Harshman and Lundy [14]. Contrived examples are given by ten Berge, Kiers, and De Leeuw [37] and Paatero [30]. The latter has also constructed sequences of CP updates with three and four diverging components.

Kruskal, Harshman, and Lundy [24] have argued that diverging CP components occur due to the fact that the array $\underline{\mathbf{X}}$ has no best rank-R approximation. They reason that every sequence of CP updates of which the objective value is approaching the infimum of the CP objective function must fail to converge and displays a pattern of diverging CP components. Stegeman [34] confirms this statement (under some regularity condition) for generic $p \times p \times 2$ arrays of rank $p+1$ with $R = p$. Stegeman [35] confirms the statement of [24] for generic $3 \times 3 \times p$ arrays with symmetric slices of rank $p+1$ with $R = p$, $p = 4, 5$, for generic $3 \times 3 \times 5$ arrays of rank 6 with $R = 5$, and for generic $8 \times 4 \times 3$ arrays of rank 9 with $R = 8$.

For given $d_1, d_2, d_3 \geq 2$, let

$$\mathcal{S}_R = \{\underline{\mathbf{Y}} \in \mathbb{R}^{d_1 \times d_2 \times d_3} : \operatorname{rank}_\circ(\underline{\mathbf{Y}}) \leq R\}\,, \tag{1.8}$$

and let $\overline{\mathcal{S}}_R$ denote its closure. We assume that $\operatorname{rank}_\circ(\underline{\mathbf{X}}) > R$. Hence, if $\underline{\mathbf{X}}$ has a best rank-R approximation, it will be a boundary point of $\mathcal{S}_R$. So far, all but one

mathematically analyzed case of diverging CP components (i.e., ten Berge, Kiers, and De Leeuw [37], Paatero [30], Stegeman [34, 35], and the results in the present paper) are due to the fact that the sequence $\underline{\mathbf{Y}}^{(n)} \in \mathcal{S}_R$ of CP updates converges to a boundary point $\underline{\widetilde{\mathbf{X}}}$ of $\mathcal{S}_R$ with $\text{rank}_\circ(\underline{\widetilde{\mathbf{X}}}) > R$, i.e., $\underline{\widetilde{\mathbf{X}}} \in \overline{\mathcal{S}}_R \backslash \mathcal{S}_R$, where $\underline{\widetilde{\mathbf{X}}}$ is a best approximation of $\underline{\mathbf{X}}$ from $\overline{\mathcal{S}}_R$. In these cases, the phenomenon of diverging CP components can be formalized as follows. There exist disjoint index sets $I_1, \ldots, I_m \subset \{1, \ldots, R\}$ such that as $\underline{\mathbf{Y}}^{(n)} \to \underline{\widetilde{\mathbf{X}}}$,

$$|\omega_r^{(n)}| \to \infty \qquad \text{for all } r \in I_j\,, \quad j = 1, \ldots, m\,, \tag{1.9}$$

$$\text{while} \quad \left\| \sum_{r \in I_j} \omega_r^{(n)} \, (\mathbf{a}_r^{(n)} \circ \mathbf{b}_r^{(n)} \circ \mathbf{c}_r^{(n)}) \right\| \qquad \text{is bounded}\,, \quad j = 1, \ldots, m\,. \tag{1.10}$$

For two diverging CP components, we have $m = 1$ and $\text{card}(I_1) = 2$. For three diverging CP components, we have $m = 1$ and $\text{card}(I_1) = 3$. For two groups of diverging CP components we have $m = 2$, et cetera. For the case of generic $p \times q \times 2$ arrays it will be shown in section 3 how the rank of $\underline{\widetilde{\mathbf{X}}}$ is related to the number of groups m and the number of diverging CP components in each group.

Note that we do not consider cases where $\text{rank}_\circ(\underline{\mathbf{X}}) = R$ and its CP decomposition resembles a case of diverging CP components, examples of which can be found in Mitchell and Burdick [28] and Paatero [30].

If $\underline{\mathbf{X}}$ does not have a best rank-R approximation, this implies that all best approximations $\underline{\widetilde{\mathbf{X}}}$ of $\underline{\mathbf{X}}$ from $\overline{\mathcal{S}}_R$ have at least rank $R + 1$. In the cases analyzed so far, any sequence $\underline{\mathbf{Y}}^{(n)}$ of CP updates converging to $\underline{\widetilde{\mathbf{X}}}$ has been shown (under some regularity conditions) to exhibit diverging CP components in this situation. Hence, modified CP algorithms designed to avoid diverging CP components (e.g., Rayens and Mitchell [31] and Cao et al. [8]) are no remedy here.

As mentioned above, there is one known case where $\underline{\mathbf{X}}$ has a best rank-R approximation, but diverging CP components may still occur. This is the case of $3 \times 3 \times 5$ arrays of rank 6 and $R = 5$. Here, a best rank-5 approximation $\underline{\widetilde{\mathbf{X}}}$ of $\underline{\mathbf{X}}$ may exist while sequences $\underline{\mathbf{Y}}^{(n)}$ of CP updates converging to $\underline{\widetilde{\mathbf{X}}}$ sometimes show diverging CP components and sometimes do not; see Stegeman [35]. This is due to the partial uniqueness of the CP decomposition of $\underline{\mathbf{Y}}^{(n)}$; see ten Berge [40].

Diverging CP components are a problem in the analysis of three-way arrays, since the obtained CP solution is hardly interpretable. Diverging CP components can be avoided by imposing orthogonality constraints on the components matrices (see Harshman and Lundy [14]) but this will come with some loss of fit. Lim [27] shows that for nonnegative $\underline{\mathbf{X}}$ and nonnegative component matrices there always exists an optimal CP solution and diverging CP components do not occur.

1.4. Matrix pencils and the rank of $p \times q \times 2$ arrays. A matrix pencil $\mathbf{X}_1 + \lambda \mathbf{X}_2$ consists of two matrices $\mathbf{X}_1$ and $\mathbf{X}_2$ with elements in a field $\mathcal{F}$ and a scalar λ. A matrix pencil is called *regular* if both $\mathbf{X}_1$ and $\mathbf{X}_2$ are square matrices and there exists an λ such that $\det(\mathbf{X}_1 + \lambda \mathbf{X}_2) \neq 0$. In all other cases, the pencil is called *singular*. For regular matrix pencils, equivalence results and a canonical form were established by Weierstrass [42]. The corresponding theory for singular pencils was developed by Kronecker [20]. For an overview of matrix pencil theory we refer the reader to Gantmacher [12, Chapter XII].

Ja' Ja' [19] has extended Kronecker's [20] equivalence results for $p \times q$ matrix pencils to $p \times q \times 2$ arrays. The same author [18] obtained results on the multiplicative complexity of computing two bilinear forms by considering the Kronecker canonical form of the associated matrix pencil and gave a complete characterization of the rank of the associated $p \times q \times 2$ array. In particular, Ja' Ja' [18] showed that for $p \geq q$,

$$\underline{\mathbf{X}} \in \mathbb{R}^{p \times q \times 2} \quad \Longrightarrow \quad \text{rank}_\circ(\underline{\mathbf{X}}) \leq q + \min(q, \text{floor}(p/2)) \,, \tag{1.11}$$

where the rank is over the real field and floor(x) denotes the largest integer smaller than or equal to x; see also Kruskal [23]. The upper bound (1.11) is sharp, i.e., there exist $p \times q \times 2$ arrays with rank equal to the upper bound. For later use, we state the following result, also due to Ja' Ja' [18].

PROPOSITION 1.2. *Let* $\underline{\mathbf{X}} \in \mathbb{R}^{p \times p \times 2}$ *with* $p \times p$ *slices* $\mathbf{X}_i$, $i = 1, 2$. *Suppose* $det(\mathbf{X}_1) \neq 0$ *and* $\mathbf{X}_2\mathbf{X}_1^{-1}$ *has* p *real eigenvalues. Let the Jordan normal form* (*see Gantmacher* [12, Chapter VI]) *of* $\mathbf{X}_2\mathbf{X}_1^{-1}$ *be given by* $\text{diag}(J_{n_1}(\lambda_1), \ldots, J_{n_r}(\lambda_r))$, *where* $J_{n_j}(\lambda_j)$ *denotes an* $n_j \times n_j$ *Jordan block with diagonal elements equal to* λ_j. *Then*

$$\text{rank}_\circ(\underline{\mathbf{X}}) = p + k \,, \tag{1.12}$$

where the rank is over the real field and k *is the number of Jordan blocks* $J_{n_j}(\lambda_j)$ *with* $n_j > 1$.

For a $p \times p$ matrix $\mathbf{Z}$ with eigenvalues $\lambda_1, \ldots, \lambda_r$, we define the *algebraic multiplicity* of λ_j as the multiplicity of λ_j as root of the characteristic polynomial $\det(\mathbf{Z} - \lambda \mathbf{I}_p)$, and the *geometric multiplicity* of λ_j as the maximum number of linearly independent eigenvectors of $\mathbf{Z}$ associated with λ_j (i.e., the dimensionality of the eigenspace of λ_j). Recall that for $\mathbf{Z} = \text{diag}(J_{n_1}(\lambda_1), \ldots, J_{n_r}(\lambda_r))$, the eigenvalues are $\lambda_1, \ldots, \lambda_r$ (not necessarily distinct), and each Jordan block $J_{n_j}(\lambda_j)$ adds n_j to the algebraic multiplicity of λ_j and 1 to the geometric multiplicity of λ_j. This establishes a relation between the eigenvalues of $\mathbf{X}_2\mathbf{X}_1^{-1}$ and the rank of the array $\underline{\mathbf{X}}$ in Proposition 1.2. In particular, if $\mathbf{X}_2\mathbf{X}_1^{-1}$ has p real eigenvalues and is diagonalizable, then $\text{rank}_\circ(\underline{\mathbf{X}}) = p$ (see also Ten Berge [38]). Ja' Ja' [18] also showed that if $\mathbf{X}_2\mathbf{X}_1^{-1}$ has at least one pair of complex eigenvalues, then $\text{rank}_\circ(\underline{\mathbf{X}}) \geq p + 1$ (see also [38]).

For generic $p \times q \times 2$ arrays, Ten Berge and Kiers [39] showed that, for $p > q$, the rank of $\underline{\mathbf{X}}$ is equal to $\min(p, 2q)$ almost everywhere, i.e., $\text{rank}(\underline{\mathbf{X}}) \neq \min(p, 2q)$ on a set of zero volume in $\mathbb{R}^{p \times q \times 2}$. We call this rank value the *typical rank*. The same authors show that for $p = q$, the typical rank of $\underline{\mathbf{X}}$ is two-valued, namely $\{p, p+1\}$, where the sets of both rank values have positive volume. Notice that for $p \times q \times 2$ arrays the set $\mathcal{S}_R$ of arrays with rank less than or equal to R has dimensionality $2pq$ if R is larger than or equal to the typical rank. If R is smaller than the typical rank, then $\mathcal{S}_R$ has dimensionality lower than $2pq$. Analogously, if R is larger than the typical rank, then the set $\mathcal{S}_R^c = \mathbb{R}^{p \times q \times 2} \backslash \mathcal{S}_R$ has dimensionality lower than $2pq$.

Notice that if $p \geq 2q$, then both the typical rank and the maximum rank (1.11) are equal to $2q$. If $2q > p > q$, then the typical rank equals p, while the maximum rank equals

$$q + \text{floor}(p/2) \geq ((p+1)/2) + \text{floor}(p/2) \geq p \,. \tag{1.13}$$

Bini [4] has studied the border rank of so-called *nondegenerate* $p \times q \times 2$ arrays, where a 3-way array $\underline{\mathbf{X}} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ is called nondegenerate if $\text{rank}_i(\underline{\mathbf{X}}) = d_i$ for $i = 1, 2, 3$. The use of the term *nondegenerate* here is a bit strange, since generic $p \times q \times 2$ arrays

are nondegenerate only if $p \leq 2q$ and $q \leq 2p$. The following result is based on the results of Ja' Ja' [18] and is due to Bini [4].

PROPOSITION 1.3. *Let* $\underline{\mathbf{X}} \in \mathbb{R}^{p \times q \times 2}$ *such that* $rank_1(\underline{\mathbf{X}}) = p$, $rank_2(\underline{\mathbf{X}}) = q$, *and* $rank_3(\underline{\mathbf{X}}) = 2$. *Let* $\mathbf{X}_1$ *and* $\mathbf{X}_2$ *be the* $p \times q$ *slices of* $\underline{\mathbf{X}}$.

(i) *Let* $p = q$ *and* $det(\mathbf{X}_1) \neq 0$. *Then* $rank_B(\underline{\mathbf{X}}) = p$ *if and only if* $\mathbf{X}_2\mathbf{X}_1^{-1}$ *has* p *real eigenvalues. If* $\mathbf{X}_2\mathbf{X}_1^{-1}$ *has at least one pair of complex eigenvalues, then* $rank_B(\underline{\mathbf{X}}) = p + 1$.

(ii) *Let* $p = q$ *and* $det(\mathbf{X}_1) = 0$. *Then* $rank_B(\underline{\mathbf{X}}) = p$ *if* $det(\mathbf{X}_1 + \lambda\mathbf{X}_2)$ *has* p *real roots* λ. *If* $det(\mathbf{X}_1 + \lambda\mathbf{X}_2)$ *has at least one pair of complex roots, then* $rank_B(\underline{\mathbf{X}}) \in \{p, p + 1\}$.

(iii) *If* $p > q$, *then* $rank_B(\underline{\mathbf{X}}) = p$ *if* $det(\mathbf{Y}_1 + \lambda\mathbf{Y}_2)$ *has only real roots* λ, *where* $\mathbf{Y}_1 + \lambda\mathbf{Y}_2$ *is the regular pencil kernel in the Kronecker canonical form of* $\mathbf{X}_1 + \lambda\mathbf{X}_2$. *If* $det(\mathbf{Y}_1 + \lambda\mathbf{Y}_2)$ *has at least one pair of complex roots, then* $rank_B(\underline{\mathbf{X}}) \in \{p, p + 1\}$.

Next, we discuss the link between the results in the present paper and the existing results mentioned above. In this paper, we consider the low-rank approximation of generic real-valued $p \times q \times 2$ arrays, where we assume $p \geq q$ without loss of generality. Such an array has typical rank $\min(p, 2q)$ and we show whether or not it has a best rank-R approximation, with $R < \min(p, 2q)$. If such a generic array is nondegenerate in the sense of Bini [4], then it has typical rank $\min(p, 2q) = p$ and $\text{rank}_\circ(\underline{\mathbf{X}}) \geq p$; see (1.4). If $\text{rank}_\circ(\underline{\mathbf{X}}) = p$, then it follows from (1.4) that also $\text{rank}_B(\underline{\mathbf{X}}) = p$. The case $\text{rank}_B(\underline{\mathbf{X}}) = p+1$ is possible only for arrays $\underline{\mathbf{X}}$ with rank larger than the typical rank p, and such arrays are not generic. Hence, our results are not covered by Proposition 1.3.

To obtain our results, we study the boundary of $\mathcal{S}_R$, the set of $p \times q \times 2$ arrays with rank at most R, and distinguish boundary arrays lying in $\mathcal{S}_R$ from those with rank larger than R. Clearly, a boundary array $\widetilde{\underline{\mathbf{X}}}$ of $\mathcal{S}_R$ has border rank at most R and is not a nondegenerate array in the sense of Bini [4] since $\text{rank}_1(\widetilde{\underline{\mathbf{X}}}) \leq R < p$; see (1.4). However, we transform the set $\mathcal{S}_R$ to a subset of the smaller space of $R \times R \times 2$ arrays if $R \leq q$, or $R \times q \times 2$ arrays if $R > q$, and Propositions 1.2 and 1.3 apply to individual boundary arrays in this smaller space. But to answer the question whether a generic $p \times q \times 2$ array has a best rank-R approximation almost everywhere, or on a set of positive volume, or on a set of zero volume, we need dimensionality arguments to establish which rank values have positive volume on the boundary of $\mathcal{S}_R$. Moreover, we do not use the Kronecker canonical form in (iii) of Proposition 1.3 to obtain our results.

This paper is organized as follows. We present our results on the existence of a best rank-R approximation to generic $p \times q \times 2$ arrays in section 2. In section 3 we show (under some regularity condition) that if there is no best rank-R approximation, then, in the CP model, any sequence $\underline{\mathbf{Y}}^{(n)} \in \mathcal{S}_R$ converging to an optimal boundary point $\widetilde{\underline{\mathbf{X}}}$ of $\mathcal{S}_R$ (with $\text{rank}_\circ(\widetilde{\underline{\mathbf{X}}}) > R$) will exhibit diverging CP components as defined by (1.9) and (1.10). Moreover, we show that there is a direct relation between the rank of $\widetilde{\underline{\mathbf{X}}}$ and the number of groups of diverging CP components. This extends Stegeman [34] who considered rank-p approximations to generic $p \times p \times 2$ arrays of rank $p + 1$. In section 4 we illustrate our results by means of calculating rank-R approximations to random $p \times q \times 2$ arrays for a variety of values for p, q, and R. Finally, section 5 contains a discussion on the presented results.

2. Low-rank approximation of generic $p \times q \times 2$ arrays. For generic $\underline{\mathbf{X}} \in \mathbb{R}^{p \times q \times 2}$, we consider the problem

$$\begin{aligned} &\text{Minimize} \|\underline{\mathbf{X}} - \underline{\mathbf{Y}}\| \\ &\text{subject to } \underline{\mathbf{Y}} \in \mathcal{S}_R\,, \end{aligned} \tag{2.1}$$

where $\mathcal{S}_R$ is the set of real-valued $p \times q \times 2$ arrays of rank at most R (see (1.8)) and the rank is taken over the real field. We also consider the related problem

$$\begin{aligned} &\text{Minimize} \|\underline{\mathbf{X}} - \underline{\mathbf{Y}}\| \\ &\text{subject to } \underline{\mathbf{Y}} \in \overline{\mathcal{S}}_R\,, \end{aligned} \tag{2.2}$$

where $\overline{\mathcal{S}}_R$ is the closure of $\mathcal{S}_R$, i.e., the union of $\mathcal{S}_R$ and all its boundary points. Suppose $\text{rank}_\circ(\underline{\mathbf{X}}) > R$. There holds that any optimal solution of problem (2.2) is a boundary point of $\mathcal{S}_R$. Indeed, from any interior point $\underline{\mathbf{Y}}$ of $\mathcal{S}_R$ a line to $\underline{\mathbf{X}}$ can be drawn which intersects with the boundary of $\mathcal{S}_R$. Suppose the intersection occurs at boundary point $\widetilde{\underline{\mathbf{X}}}$. Then $\widetilde{\underline{\mathbf{X}}}$ has a lower objective value (i.e., is closer to $\underline{\mathbf{X}}$) than the interior point $\underline{\mathbf{Y}}$. Hence, problem (2.1) has an optimal solution, or, equivalently, $\underline{\mathbf{X}}$ has a best rank-R approximation, if there exists a boundary point $\widetilde{\underline{\mathbf{X}}} \in \mathcal{S}_R$ which is an optimal solution of problem (2.2). Clearly, this always holds if $\mathcal{S}_R$ is a closed set. However, De Silva and Lim [10] have shown that $\mathcal{S}_R$ is closed only for $R = 1$. Hence, an investigation of the boundary points of $\mathcal{S}_R$ is necessary to ascertain whether a best rank-R approximation exists almost everywhere, or on a set of positive volume, or on a set of zero volume.

We consider all possible combinations of p, q, and R, where, without loss of generality, we assume that $p \geq q$. As mentioned in section 1, we transform the set $\mathcal{S}_R$ to a subset of the smaller space of $R \times R \times 2$ arrays if $R \leq q$ or $R \times q \times 2$ arrays if $R > q$ and use the results of Stegeman [34] and Proposition 1.2 to characterize the boundary points of $\mathcal{S}_R$; in particular, whether they are in $\mathcal{S}_R$ or in $\mathcal{S}_R^c$.

Unfortunately, our results are not complete. That is, they rely on conjectures relating the dimensionality of parts of the boundary of $\mathcal{S}_R$ to the existence of optimal solutions of problem (2.1). Table 2.1 gives a summary of our results. Below, we consider all cases in Table 2.1 and state explicitly whether we use a conjecture. Except for cases 1, 4, and 6, the statements on the existence of a best rank-R approximation are (partly) based on conjectures.

Cases 1, 4, and 6. In these cases, R is larger than or equal to the typical rank of $\underline{\mathbf{X}}$, i.e., $\underline{\mathbf{X}}$ itself lies in $\mathcal{S}_R$ almost everywhere. Hence, the best rank-R approximation of $\underline{\mathbf{X}}$ is $\underline{\mathbf{X}}$ itself.

Cases 2, 3, 5, 7, 8, and 9. For these cases, there holds $\underline{\mathbf{X}} \notin \mathcal{S}_R$ almost everywhere. As discussed above, we need to characterize the boundary of $\mathcal{S}_R$. We define the following subsets of $\mathbb{R}^{p \times q \times 2}$. Let

$$\mathcal{W}_R = \{\underline{\mathbf{Y}} \in \mathbb{R}^{p \times q \times 2} : \text{rank}[\mathbf{Y}_1 | \mathbf{Y}_2] \leq R\} = \{\underline{\mathbf{Y}} \in \mathbb{R}^{p \times q \times 2} : \text{rank}_1(\underline{\mathbf{Y}}) \leq R\}\,, \tag{2.3}$$

where $\mathbf{Y}_1$ and $\mathbf{Y}_2$ denote the $p \times q$ slices of $\underline{\mathbf{Y}}$, and let

$$\mathcal{V}_{R+1} = \{\underline{\mathbf{Y}} \in \mathcal{W}_R : \text{rank}_\circ(\underline{\mathbf{Y}}) \geq R + 1\} = \mathcal{W}_R \cap \mathcal{S}_R^c\,. \tag{2.4}$$

We need the following lemma.

LEMMA 2.1. *We have the following results.*

(i) *The set $\mathcal{W}_R$ is closed and the boundary of $\mathcal{W}_R$ is the set $\mathcal{W}_R$ itself.*

TABLE 2.1
Results (cases 1, 4, and 6) and conjectures (cases 2, 3, 5, 7, 8, and 9) on the existence of a best rank-R approximation to generic $p \times q \times 2$ arrays. Here, $p \geq q \geq 2$ and $R \geq 2$.

Case	$\underline{\mathbf{X}} \in \mathbb{R}^{p \times q \times 2}$	$\text{rank}_\circ(\underline{\mathbf{X}})$	R	Best rank-R approx. exists?
1	$p = q$	$p+1$	$R \geq p+1$	almost everywhere
2	$p = q$	$p+1$	$R = p$	zero volume
3	$p = q$	$p+1$	$R < p$	positive volume
4	$p = q$	p	$R \geq p$	almost everywhere
5	$p = q$	p	$R < p$	positive volume
6	$p > q$	$\min(p, 2q)$	$R \geq \min(p, 2q)$	almost everywhere
7	$p > q$	$\min(p, 2q)$	$\min(p, 2q) > R > q$	almost everywhere
8	$p > q$	$\min(p, 2q)$	$R = q$	positive volume
9	$p > q$	$\min(p, 2q)$	$R < q$	positive volume

(ii) *There holds*

$$\mathcal{W}_R = \mathcal{S}_R \cup \mathcal{V}_{R+1} \qquad \text{and} \qquad \mathcal{S}_R \cap \mathcal{V}_{R+1} = \emptyset\,. \tag{2.5}$$

(iii) *In case 2, there holds $\mathcal{W}_R = \mathbb{R}^{p \times q \times 2}$.*

Proof. Statement (i) follows from the fact that any matrix of rank at most R can be approximated arbitrarily well by rank-$(R+1)$ matrices.

Next, we prove (ii). From (1.4) it follows that $\mathcal{S}_R \subseteq \mathcal{W}_R$. Since $\mathcal{V}_{R+1} = \mathcal{W}_R \cap \mathcal{S}_R^c$, we have $\mathcal{S}_R \cap \mathcal{V}_{R+1} = \emptyset$ by definition. Hence, $\mathcal{W}_R = (\mathcal{W}_R \cap \mathcal{S}_R) \cup (\mathcal{W}_R \cap \mathcal{S}_R^c) = \mathcal{S}_R \cup \mathcal{V}_{R+1}$.

In case 2, we have $R = p = q$, which implies that $[\mathbf{Y}_1 | \, \mathbf{Y}_2]$ is a matrix of order $p \times 2p$. Obviously, this matrix always has rank less than or equal to p. Hence, $\mathcal{W}_R = \mathbb{R}^{p \times q \times 2}$ in case 2. This proves (iii). ∎

Next, we consider the boundary points of $\mathcal{S}_R$. The complement of $\mathcal{S}_R$ is equal to

$$\mathcal{S}_R^c = \mathcal{W}_R^c \cup \mathcal{V}_{R+1} \qquad \text{with} \qquad \mathcal{W}_R^c \cap \mathcal{V}_{R+1} = \emptyset\,. \tag{2.6}$$

Hence, the boundary of $\mathcal{S}_R$ consists of the boundary between $\mathcal{S}_R$ and $\mathcal{V}_{R+1}$ and the boundary between $\mathcal{S}_R$ and $\mathcal{W}_R^c$. Note that these two boundaries may have a nonempty intersection. We denote the boundary of $\mathcal{S}_R$ as $\partial\mathcal{S}_R$ and partition it into the following two sets. Let

$$\mathcal{U}_R^{(1)} = \partial\mathcal{S}_R \cap \partial\mathcal{V}_{R+1} \qquad \text{and} \qquad \mathcal{U}_R^{(2)} = \partial\mathcal{S}_R \cap (\partial\mathcal{V}_{R+1})^c\,. \tag{2.7}$$

Hence, $\mathcal{U}_R^{(1)}$ consists of all points on the boundary between $\mathcal{S}_R$ and $\mathcal{V}_{R+1}$, i.e., all points which can be approximated arbitrarily well from both $\mathcal{S}_R$ and $\mathcal{V}_{R+1}$.

From Lemma 2.1 it follows that

$$\partial\mathcal{S}_R \subseteq \mathcal{W}_R = \mathcal{S}_R \cup \mathcal{V}_{R+1}\,. \tag{2.8}$$

The following lemma states that $\mathcal{U}_R^{(2)}$ is either the empty set or is a subset of $\mathcal{S}_R$.

LEMMA 2.2. *We have the following results.*

(i) *In case 2, there holds $\mathcal{U}_R^{(2)} = \emptyset$.*

(ii) *In cases* 3, 5, 7, 8, *and* 9, *there holds* $\mathcal{U}_R^{(2)} \subseteq \mathcal{S}_R$.

Proof. From (2.8) it follows that $\mathcal{U}_R^{(1)} \subseteq \mathcal{W}_R$ and $\mathcal{U}_R^{(2)} \subseteq \mathcal{W}_R$. In case 2, we have $\mathcal{W}_R = \mathbb{R}^{p\times q\times 2}$ (see Lemma 2.1) and, hence, $\mathcal{V}_{R+1} = \mathcal{S}_R^c$. This implies that $\mathcal{U}_R^{(1)}$ consists of all points which can be approximated arbitrarily well from $\mathcal{S}_R$ and $\mathcal{S}_R^c$. Since these are all boundary points of $\mathcal{S}_R$, it follows that $\mathcal{U}_R^{(2)} = \emptyset$. This proves (i).

Next, consider cases 3, 5, 7, 8, and 9. Here, $\mathcal{W}_R^c \neq \emptyset$. Suppose $\underline{\mathbf{Y}} \in \mathcal{V}_{R+1}$ and $\underline{\mathbf{Y}} \in \partial\mathcal{S}_R$. Since $\mathcal{S}_R \cap \mathcal{V}_{R+1} = \emptyset$, it follows that $\underline{\mathbf{Y}} \in \partial\mathcal{V}_{R+1}$ and, hence, $\underline{\mathbf{Y}} \in \mathcal{U}_R^{(1)}$. This implies $\mathcal{V}_{R+1} \cap \mathcal{U}_R^{(2)} = \emptyset$. Moreover, from $\mathcal{U}_R^{(2)} \subseteq \mathcal{W}_R = \mathcal{S}_R \cup \mathcal{V}_{R+1}$, we obtain $\mathcal{U}_R^{(2)} \subseteq \mathcal{S}_R$. This proves (ii). ∎

In the remaining part of this section, we consider the part $\mathcal{U}_R^{(1)}$ of the boundary of $\mathcal{S}_R$. For each of the cases 2, 3, 5, 7, 8 and 9 in Table 2.1, we argue that the nonexistence of a best rank-R approximation is due to the fact that $\mathcal{U}_R^{(1)} \not\subseteq \mathcal{S}_R$.

Case 2. We have $p = q = R$, $\mathcal{W}_R = \mathbb{R}^{p\times q\times 2} = \mathcal{S}_R \cup \mathcal{V}_{R+1}$, $\mathcal{V}_{R+1} = \mathcal{S}_R^c$, $\mathcal{U}_R^{(2)} = \emptyset$, and $\underline{\mathbf{X}} \in \mathcal{S}_R^c$. The typical rank of $p \times p \times 2$ arrays is equal to $\{p,\, p+1\}$, where the sets of both rank values have positive volume. This implies that the sets $\mathcal{S}_R$ and $\mathcal{V}_{R+1}$ have equal dimensionality $2p^2$. We partition $\partial\mathcal{S}_R = \mathcal{U}_R^{(1)}$ into the following three sets. Let

$$\mathcal{U}_R^{(11)} = \{\underline{\mathbf{Y}} \in \mathcal{U}_R^{(1)} \text{ with } \mathbf{Y}_1 \text{ nonsingular and } \mathbf{Y}_2\mathbf{Y}_1^{-1} \text{ diagonalizable}\}, \tag{2.9}$$

$$\mathcal{U}_R^{(12)} = \{\underline{\mathbf{Y}} \in \mathcal{U}_R^{(1)} \text{ with } \mathbf{Y}_1 \text{ nonsingular and } \mathbf{Y}_2\mathbf{Y}_1^{-1} \text{ not diagonalizable}\}, \tag{2.10}$$

$$\mathcal{U}_R^{(13)} = \{\underline{\mathbf{Y}} \in \mathcal{U}_R^{(1)} \text{ with } \mathbf{Y}_1 \text{ singular}\}. \tag{2.11}$$

In Stegeman [34], it is shown that $\underline{\mathbf{Y}} \in \mathcal{U}_R^{(11)} \cup \mathcal{U}_R^{(12)}$ if and only if $\mathbf{Y}_2\mathbf{Y}_1^{-1}$ has p real eigenvalues which are not all distinct. Although its proof is entirely different, this result is closely related to Propositions 1.2 and 1.3(i). From Proposition 1.2 it follows that $\mathcal{U}_R^{(11)} \subset \mathcal{S}_R$ and $\mathcal{U}_R^{(12)} \subset \mathcal{S}_R^c$. Stegeman [34] also shows that the dimensionality of $\mathcal{U}_R^{(11)}$ is lower that the dimensionality of $\mathcal{U}_R^{(12)}$.

Any array in $\mathcal{U}_R^{(13)}$ can be approximated arbitrarily well by arrays in $\mathcal{U}_R^{(11)} \cup \mathcal{U}_R^{(12)}$. The reverse, however, is not true. This implies that $\mathcal{U}_R^{(13)}$ has lower dimensionality than $\mathcal{U}_R^{(11)} \cup \mathcal{U}_R^{(12)}$. Combined with the reasoning above, this implies that the subset of the boundary of $\mathcal{S}_R$ with the highest dimensionality is $\mathcal{U}_R^{(12)}$. Since these boundary points have rank larger than R, we conjecture that for a generic $\underline{\mathbf{X}} \in \mathbb{R}^{p\times p\times 2}$ array an optimal solution $\underline{\widetilde{\mathbf{X}}}$ of problem (2.2) has rank larger than R almost everywhere. Hence, we conjecture that problem (2.1) does not have an optimal solution almost everywhere, and $\underline{\mathbf{X}}$ does not have a best rank-R approximation almost everywhere.

Case 7. We have $p > q$ and $\min(p, 2q) > R > q$. From the definition of $\mathcal{W}_R$ in (2.3) it follows that $\underline{\mathbf{Y}} \in \mathcal{W}_R$ if and only if there exists a nonsingular matrix $\mathbf{S}$ such that

$$\mathbf{S}\,\mathbf{Y}_1 = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{O} \end{bmatrix} \qquad \text{and} \qquad \mathbf{S}\,\mathbf{Y}_2 = \begin{bmatrix} \mathbf{H}_2 \\ \mathbf{O} \end{bmatrix}, \tag{2.12}$$

where $\mathbf{H}_1$ and $\mathbf{H}_2$ are $R \times q$ matrices and $\mathbf{O}$ is the $(p-R) \times q$ all-zero matrix. By $\underline{\mathbf{H}}$ we denote the $R \times q \times 2$ array with slices $\mathbf{H}_1$ and $\mathbf{H}_2$. The transformation $\mathbf{S}$ in (2.12)

is rank preserving, i.e., $\text{rank}_\circ(\underline{\mathbf{Y}}) = \text{rank}_\circ(\underline{\mathbf{H}})$. Hence, it follows from (2.4) and (2.5) that

$$\mathcal{S}_R = \{\underline{\mathbf{Y}} \in \mathcal{W}_R : \ \text{rank}_\circ(\underline{\mathbf{H}}) \leq R \ \text{ for } \underline{\mathbf{H}} \text{ in (2.12)}\} \tag{2.13}$$

and

$$\mathcal{V}_{R+1} = \{\underline{\mathbf{Y}} \in \mathcal{W}_R : \ \text{rank}_\circ(\underline{\mathbf{H}}) \geq R+1 \ \text{ for } \underline{\mathbf{H}} \text{ in (2.12)}\}\,. \tag{2.14}$$

It is important to note that the set of arrays $\underline{\mathbf{H}}$ which can be obtained in (2.12) has full dimensionality $2qR$. The typical rank of $R \times q \times 2$ arrays (with $2q > R > q$) equals R. When the maximal rank of $R \times q \times 2$ arrays equals their typical rank, we have $\mathcal{V}_{R+1} = \emptyset$, and $\mathcal{S}_R = \mathcal{W}_R$ is a closed set. This is the case when $R = 4$ and $q = 3$, e.g., see (1.11). Since $\mathcal{S}_R$ is closed, a best rank-R approximation to $\underline{\mathbf{X}}$ exists almost everywhere.

Next, suppose the maximal rank of $R \times q \times 2$ arrays is larger than their typical rank R, i.e., $\mathcal{V}_{R+1} \neq \emptyset$. This is the case when $R = 5$ and $q = 3$, e.g., see (1.11). For the arrays in $\mathcal{V}_{R+1}$, the array $\underline{\mathbf{H}}$ has a *nontypical* rank value, i.e., larger than or equal to $R+1$. This implies that $\mathcal{V}_{R+1}$ has lower dimensionality than the set $\mathcal{S}_R$. From (2.4) and (2.5) it then follows that any $\underline{\mathbf{Y}} \in \mathcal{V}_{R+1}$ can be approximated arbitrarily closely by arrays in $\mathcal{S}_R$ and it holds that $\mathcal{U}_R^{(1)} = \mathcal{V}_{R+1}$; see (2.7). Moreover, any array in $\mathcal{S}_R$ can be approximated arbitrarily well from $\mathcal{W}_R^c$ and, hence, $\mathcal{U}_R^{(2)} = \mathcal{S}_R$. This implies that

$$\partial\mathcal{S}_R = \mathcal{V}_{R+1} \cup \mathcal{S}_R = \mathcal{W}_R = \overline{\mathcal{S}}_R\,. \tag{2.15}$$

Since $\mathcal{V}_{R+1}$ has lower dimensionality than $\mathcal{S}_R$, we conjecture that an optimal solution $\widetilde{\underline{\mathbf{X}}}$ of problem (2.2) lies in $\mathcal{S}_R$ almost everywhere. Hence, we conjecture that $\underline{\mathbf{X}}$ has a best rank-R approximation almost everywhere.

Case 8. We have $p > q$ and $R = q$. As in case 7, there holds $\underline{\mathbf{Y}} \in \mathcal{W}_R$ if and only if a nonsingular $\mathbf{S}$ exists such that (2.12) holds. The sets $\mathcal{S}_R$ and $\mathcal{V}_{R+1}$ are defined by (2.13) and (2.14), respectively. The array $\underline{\mathbf{H}}$ in (2.12) has order $R \times R \times 2$ and the typical rank of $R \times R \times 2$ arrays is equal to $\{R,\, R+1\}$, where the sets of both rank values have positive volume. As in case 2, this implies that the sets $\mathcal{S}_R$ and $\mathcal{V}_{R+1}$ have equal dimensionality.

Let $\underline{\mathbf{X}}^*$ be an optimal solution of the following problem:

$$\begin{aligned} &\text{Minimize} \|\underline{\mathbf{X}} - \underline{\mathbf{Y}}\| \\ &\text{subject to } \underline{\mathbf{Y}} \in \mathcal{W}_R\,. \end{aligned} \tag{2.16}$$

Based on Lemma 2.1, we conjecture that, for generic $\underline{\mathbf{X}}$, the set where $\underline{\mathbf{X}}^* \in \mathcal{S}_R$ and the set where $\underline{\mathbf{X}}^* \in \mathcal{V}_{R+1}$ both have positive volume. If $\underline{\mathbf{X}}^* \in \mathcal{S}_R$, then $\underline{\mathbf{X}}^*$ is an optimal solution of the problem (2.1) and, hence, $\underline{\mathbf{X}}$ has a best rank-R approximation. If all optimal solutions of problem (2.16) lie in $\mathcal{V}_{R+1}$, then $\underline{\mathbf{X}}$ may not have a best rank-R approximation. This will be explained below.

The boundary of $\mathcal{S}_R$ is partitioned into $\mathcal{U}_R^{(1)}$ and $\mathcal{U}_R^{(2)}$, where $\mathcal{U}_R^{(2)} \subseteq \mathcal{S}_R$; see (2.7) and Lemma 2.2. Analogous to case 2, we partition $\mathcal{U}_R^{(1)}$ into the following sets:

$$\mathcal{U}_R^{(11)} = \{\underline{\mathbf{Y}} \in \mathcal{U}_R^{(1)} \text{ with } \mathbf{H}_1 \text{ in (2.12) nonsingular and } \mathbf{H}_2\mathbf{H}_1^{-1} \text{ diagonalizable}\}\,, \tag{2.17}$$

$$\mathcal{U}_R^{(12)} = \{\underline{\mathbf{Y}} \in \mathcal{U}_R^{(1)} \text{ with } \mathbf{H}_1 \text{ in (2.12) nonsingular and } \mathbf{H}_2\mathbf{H}_1^{-1} \text{ not diagonalizable}\}, \tag{2.18}$$

$$\mathcal{U}_R^{(13)} = \{\underline{\mathbf{Y}} \in \mathcal{U}_R^{(1)} \text{ with } \mathbf{H}_1 \text{ in (2.12) singular}\}. \tag{2.19}$$

Note that since $\text{rank}(\mathbf{H}_1) = \text{rank}(\mathbf{S}\,\mathbf{Y}_1) = \text{rank}(\mathbf{Y}_1)$ and $\text{rank}_\circ(\underline{\mathbf{Y}}) = \text{rank}_\circ(\mathbf{H})$, combined with Proposition 1.2, it follows that the sets in (**??**)–(2.19) do not depend on the choice of $\mathbf{S}$ in (2.12). Analogous to case 2, there holds $\underline{\mathbf{Y}} \in \mathcal{U}_R^{(11)} \cup \mathcal{U}_R^{(12)}$ if and only if $\mathbf{H}_2\mathbf{H}_1^{-1}$ has R real eigenvalues which are not all distinct. Also, we have $\mathcal{U}_R^{(11)} \subset \mathcal{S}_R$ and $\mathcal{U}_R^{(12)} \subset \mathcal{S}_R^c$. Moreover, the sets $\mathcal{U}_R^{(11)}$ and $\mathcal{U}_R^{(13)}$ have lower dimensionality than $\mathcal{U}_R^{(12)}$. Hence, if all optimal solutions of problem (2.2) lie in $\mathcal{U}_R^{(1)}$, then they have rank larger than R almost everywhere on $\mathcal{U}_R^{(1)}$, and $\underline{\mathbf{X}}$ does not have a best rank-R approximation.

Above, we conjectured that, for generic $\underline{\mathbf{X}}$, an optimal solution $\underline{\mathbf{X}}^*$ of problem (2.16) lies in $\mathcal{S}_R$ on a set of positive volume. Hence, we conjecture that $\underline{\mathbf{X}}$ has a best rank-R approximation on a set of positive volume. Next, we argue that the set on which $\underline{\mathbf{X}}$ has no best rank-R approximation also has positive volume. This can be seen as follows. Let $\mathbf{Y} \in \mathcal{V}_{R+1}$ be an interior point of $\mathcal{V}_{R+1}$ on $\mathcal{W}_R$, i.e., for a small $\epsilon > 0$ we have $\mathcal{B}_\epsilon(\underline{\mathbf{Y}}) = \{\underline{\mathbf{Z}} \in \mathcal{W}_R : \ \|\underline{\mathbf{Z}} - \underline{\mathbf{Y}}\| < \epsilon\} \subset \mathcal{V}_{R+1}$. This is the case if $\mathbf{H}_2(\mathbf{H}_1)^{-1}$ has R distinct eigenvalues which are not all real, where $\mathbf{H}_i$ are as in (2.12); see Stegeman [34]. For any interior point $\underline{\mathbf{Y}}$ of $\mathcal{V}_{R+1}$, we can find a set $\mathcal{D} \subset \mathcal{W}_R^c$ close to $\mathcal{B}_\epsilon(\underline{\mathbf{Y}})$ such that $\mathcal{D}$ has positive volume and for any $\underline{\mathbf{X}} \in \mathcal{D}$ problem (2.16) will have all optimal solutions in $\mathcal{B}_\epsilon(\underline{\mathbf{Y}})$. Moreover, for $\underline{\mathbf{Y}}$ close enough to some point on $\mathcal{U}_R^{(12)}$, i.e., the boundary between $\mathcal{S}_R$ and $\mathcal{V}_{R+1}$, we conjecture that all optimal solutions of problem (2.2) will lie on the boundary $\mathcal{U}_R^{(12)}$ and have rank larger than R. Hence, we conjecture that no $\underline{\mathbf{X}} \in \mathcal{D}$ has a best rank-R approximation, where $\mathcal{D}$ has positive volume. Therefore, we conjecture that the set on which $\underline{\mathbf{X}}$ has no best rank-R approximation has positive volume.

Cases 3, 5, and 9. We have $p \geq q$ and $R < q$. Instead of $\mathcal{W}_R$, we define

$$\widetilde{\mathcal{W}}_R = \left\{\underline{\mathbf{Y}} \in \mathbb{R}^{p\times q\times 2} : \ \text{rank}[\mathbf{Y}_1|\,\mathbf{Y}_2] \leq R \text{ and rank} \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} \leq R\right\}$$

$$= \{\underline{\mathbf{Y}} \in \mathbb{R}^{p\times q\times 2} : \ \text{rank}_1(\underline{\mathbf{Y}}) \leq R \text{ and } \text{rank}_2(\underline{\mathbf{Y}}) \leq R\}. \tag{2.20}$$

Analogous to Lemma 2.1, the set $\widetilde{\mathcal{W}}_R$ is closed. It can be seen that $\underline{\mathbf{Y}} \in \widetilde{\mathcal{W}}_R$ if and only if there exist nonsingular $\mathbf{S}$ and $\mathbf{T}$ such that

$$\mathbf{S}\,\mathbf{Y}_1\,\mathbf{T} = \begin{bmatrix} \mathbf{G}_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} \qquad \text{and} \qquad \mathbf{S}\,\mathbf{Y}_2\,\mathbf{T} = \begin{bmatrix} \mathbf{G}_2 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, \tag{2.21}$$

where $\mathbf{G}_1$ and $\mathbf{G}_2$ are $R \times R$ matrices. We denote the $R \times R \times 2$ array with slices $\mathbf{G}_1$ and $\mathbf{G}_2$ by $\underline{\mathbf{G}}$. Note that the set of arrays $\underline{\mathbf{G}}$ which can be obtained by (2.21) has full dimensionality $2R^2$. Since the transformations $\mathbf{S}$ and $\mathbf{T}$ are rank preserving, we have, analogous to (2.13), that

$$\mathcal{S}_R = \{\underline{\mathbf{Y}} \in \widetilde{\mathcal{W}}_R : \ \text{rank}_\circ(\underline{\mathbf{G}}) \leq R \ \text{ for } \underline{\mathbf{G}} \text{ in (2.21)}\}. \tag{2.22}$$

We define

$$\widetilde{\mathcal{V}}_{R+1} = \{\underline{\mathbf{Y}} \in \widetilde{\mathcal{W}}_R : \text{ rank}_\circ(\underline{\mathbf{G}}) \geq R+1 \text{ for } \underline{\mathbf{G}} \text{ in (2.21)}\}. \tag{2.23}$$

Analogous to Lemma 2.1, there holds

$$\widetilde{\mathcal{W}}_R = \mathcal{S}_R \cup \widetilde{\mathcal{V}}_{R+1} \qquad \text{and} \qquad \mathcal{S}_R \cap \widetilde{\mathcal{V}}_{R+1} = \emptyset. \tag{2.24}$$

The typical rank of $R \times R \times 2$ arrays is equal to $\{R,\, R+1\}$, where both sets of rank values have positive volume. As in case 2, this implies that the sets $\mathcal{S}_R$ and $\widetilde{\mathcal{V}}_{R+1}$ have equal dimensionality. We are in the same situation as in case 8, with $\widetilde{\mathcal{W}}_R$ playing the role of $\mathcal{W}_R$, $\widetilde{\mathcal{V}}_{R+1}$ playing the role of $\mathcal{V}_{R+1}$, and $\underline{\mathbf{G}}$ playing the role of $\underline{\mathbf{H}}$. The remaining part of the explanation for cases 3, 5, and 9 is completely analogous to case 8 and is therefore omitted.

3. Diverging CP components for $p \times q \times 2$ arrays. In the previous section, we showed that if $\underline{\mathbf{X}}$ does not have a best rank-R approximation, this is due to the fact that the optimal solutions of problem (2.2) do not lie in $\mathcal{S}_R$, i.e., they have rank of at least $R+1$. From now on we assume there is one optimal solution $\widetilde{\underline{\mathbf{X}}}$ of problem (2.2). The general case is completely analogous. As a regularity condition, we assume that $\widetilde{\underline{\mathbf{X}}} \in \partial\mathcal{S}_R \cap \mathcal{S}_R^c$ lies in a subset $\mathcal{Q}_{R+1} \subset \partial\mathcal{S}_R \cap \mathcal{S}_R^c$ such that the dimensionality of $(\partial\mathcal{S}_R \cap \mathcal{S}_R^c) \backslash \mathcal{Q}_{R+1}$ is lower than the dimensionality of $\mathcal{Q}_{R+1}$ itself. In each of cases 2, 3, 5, 8, and 9 of Table 2.1, the set $\mathcal{Q}_{R+1}$ will be specified, and we show that if a sequence of CP updates $\underline{\mathbf{Y}}^{(n)}$ converges to $\widetilde{\underline{\mathbf{X}}}$, then $\underline{\mathbf{Y}}^{(n)}$ will exhibit diverging CP components as defined by (1.9) and (1.10). This implies that in these cases of Table 2.1, we conjecture that diverging CP components occur almost everywhere (case 2) or on a set of positive volume (cases 3, 5, 8, and 9). In Table 3.1 these conjectures are stated explicitly.

TABLE 3.1
Conjectures on the occurrence of diverging CP components when calculating a best rank-R approximation to generic $p \times q \times 2$ arrays. Here, $p \geq q \geq 2$ and $R \geq 2$.

Case	$\underline{\mathbf{X}} \in \mathbb{R}^{p\times q\times 2}$	$\text{rank}_\circ(\underline{\mathbf{X}})$	R	Diverging CP components?
2	$p = q$	$p+1$	$R = p$	almost everywhere
3	$p = q$	$p+1$	$R < p$	positive volume
5	$p = q$	p	$R < p$	positive volume
7	$p > q$	$\min(p, 2q)$	$\min(p, 2q) > R > q$	zero volume
8	$p > q$	$\min(p, 2q)$	$R = q$	positive volume
9	$p > q$	$\min(p, 2q)$	$R < q$	positive volume

Case 2. We have $p = q = R$ and $\text{rank}_\circ(\underline{\mathbf{X}}) = p+1$. We assume that $\widetilde{\underline{\mathbf{X}}}$ lies in $\mathcal{U}_R^{(12)}$, which is the set $\mathcal{Q}_{R+1}$ in this case. The set $\mathcal{U}_R^{(12)}$ is defined by (2.10). This implies that $\widetilde{\mathbf{X}}_2\widetilde{\mathbf{X}}_1^{-1}$ has p real eigenvalues and is not diagonalizable; see Stegeman [34]. From Proposition 1.2 it follows that $\text{rank}_\circ(\widetilde{\underline{\mathbf{X}}}) = p+k$, where k is the number of eigenvalues of $\widetilde{\mathbf{X}}_2\widetilde{\mathbf{X}}_1^{-1}$ with algebraic multiplicity larger than the geometric multiplicity. Suppose the sequence of CP updates $\underline{\mathbf{Y}}^{(n)}$ converges to $\widetilde{\underline{\mathbf{X}}}$. Since $\underline{\mathbf{Y}}^{(n)} \in \mathcal{S}_R$ and a singular $\mathbf{Y}_1^{(n)}$ does not occur in practice, it follows from Proposition 1.2 that

$\mathbf{Y}_2^{(n)}(\mathbf{Y}_1^{(n)})^{-1}$ has p real eigenvalues and is diagonalizable. Let $\mathbf{Y}_2^{(n)}(\mathbf{Y}_1^{(n)})^{-1}$ have the eigendecomposition

$$\mathbf{Y}_2^{(n)}(\mathbf{Y}_1^{(n)})^{-1} = \mathbf{K}^{(n)}\mathbf{\Lambda}^{(n)}(\mathbf{K}^{(n)})^{-1}\,, \tag{3.1}$$

where $\mathbf{K}^{(n)}$ has columns of length 1. A rank-p decomposition (1.6) of $\underline{\mathbf{Y}}^{(n)}$ is given by

$$\mathbf{A}^{(n)} = \mathbf{K}^{(n)}\,, \qquad (\mathbf{B}^{(n)})^T = \mathbf{\Omega}_b^{(n)}\,(\mathbf{K}^{(n)})^{-1}\mathbf{Y}_1^{(n)}\,, \tag{3.2}$$

$$\mathbf{C}^{(n)} = \begin{bmatrix} 1 & \cdots & 1 \\ \lambda_1^{(n)} & \cdots & \lambda_p^{(n)} \end{bmatrix} \mathbf{\Omega}_c^{(n)}\,, \qquad \mathbf{\Omega}^{(n)} = (\mathbf{\Omega}_b^{(n)}\,\mathbf{\Omega}_c^{(n)})^{-1}\,, \tag{3.3}$$

where $\mathbf{\Omega}_b^{(n)}$ and $\mathbf{\Omega}_c^{(n)}$ are the $p \times p$ diagonal matrices such that the columns of $\mathbf{B}^{(n)}$ and $\mathbf{C}^{(n)}$, respectively, have length 1. Hence, $\mathbf{C}_1^{(n)} = \mathbf{\Omega}_c^{(n)}$ and $\mathbf{C}_2^{(n)} = \mathbf{\Lambda}^{(n)}\mathbf{\Omega}_c^{(n)}$ in (1.6). It is clear that $k_{\mathbf{A}^{(n)}} = p$, $k_{\mathbf{B}^{(n)}} = p$, and $k_{\mathbf{C}^{(n)}} = 2$ if the eigenvalues in $\mathbf{\Lambda}^{(n)}$ are distinct. In this case, Kruskal's condition (1.7) holds and the rank-p decomposition (3.2)–(3.3) of $\underline{\mathbf{Y}}^{(n)}$ is unique. Since identical eigenvalues in $\mathbf{\Lambda}^{(n)}$ do not occur in practice, we assume they are all distinct (see Stegeman [34] for identical eigenvalues).

By continuity, the matrix $\mathbf{Y}_2^{(n)}(\mathbf{Y}_1^{(n)})^{-1}$ will converge to $\widetilde{\mathbf{X}}_2\widetilde{\mathbf{X}}_1^{-1}$. Since the latter matrix does not have p linearly independent eigenvectors and the rank-p decomposition (3.2)–(3.3) is unique, it follows that $\mathbf{A}^{(n)}$ in (3.2) converges to the singular matrix of eigenvectors of $\widetilde{\mathbf{X}}_2\widetilde{\mathbf{X}}_1^{-1}$. Let $I_1, \ldots, I_m \subset \{1, \ldots, R\}$ be the disjoint index sets of linearly dependent columns of the latter matrix such that each I_j contains the linearly dependent eigenvectors associated with a different eigenvalue of $\widetilde{\mathbf{X}}_2\widetilde{\mathbf{X}}_1^{-1}$ which has algebraic multiplicity larger than the geometric multiplicity. Then the columns $I_1, \ldots, I_m$ of the matrix $(\mathbf{K}^{(n)})^{-1}$ will become arbitrarily large as $\underline{\mathbf{Y}}^{(n)} \to \underline{\widetilde{\mathbf{X}}}$. The columns I_j of $\mathbf{C}^{(n)}$ will become identical, since these correspond to identical eigenvalues of $\widetilde{\mathbf{X}}_2\widetilde{\mathbf{X}}_1^{-1}$, $j = 1, \ldots, m$. It follows that in (3.2)–(3.3), we have $|\omega_r^{(n)}| \to \infty$ for all $r \in I_j$, $j = 1, \ldots, m$. Hence, (1.9) holds.

Next, we show that (1.10) also holds. For $j \in \{1, \ldots, m\}$, we consider the contribution of components I_j to the rank-p decomposition (3.2)–(3.3) for slices $\mathbf{Y}_1^{(n)}$ and $\mathbf{Y}_2^{(n)}$ separately. For $\mathbf{Y}_1^{(n)}$, the contribution of I_j equals

$$\mathbf{P}_j^{(n)} = \mathbf{K}_j^{(n)}\,(\mathbf{K}^{(n)})_j^{-1}\mathbf{Y}_1^{(n)}\,, \tag{3.4}$$

where $\mathbf{K}_j^{(n)}$ denotes columns I_j of $\mathbf{K}^{(n)}$ and $(\mathbf{K}^{(n)})_j^{-1}$ denotes rows I_j of $(\mathbf{K}^{(n)})^{-1}$. The limit point of $\mathbf{K}^{(n)}$ has columns I_j linearly dependent, while they are linearly independent of all its other columns. Hence, for n large enough, $\mathbf{K}^{(n)}$ will have columns I_j close to linear dependence but linearly independent of all its other columns. This, together with $\|\mathbf{Y}_1^{(n)}\|$ being bounded, yields that $\|\mathbf{P}_j^{(n)}\|$ in (3.4) is bounded.

For $\mathbf{Y}_2^{(n)}$, the contribution of I_j equals

$$\mathbf{Q}_j^{(n)} = \mathbf{K}_j^{(n)}\,\mathbf{\Lambda}_j^{(n)}\,(\mathbf{K}^{(n)})_j^{-1}\mathbf{Y}_1^{(n)}\,, \tag{3.5}$$

where $\mathbf{\Lambda}_j^{(n)}$ denotes the submatrix of $\mathbf{\Lambda}^{(n)}$ containing rows I_j and columns I_j. The diagonal matrix $\mathbf{\Lambda}_j^{(n)}$ converges to $\lambda\,\mathbf{I}_z$, where λ is the eigenvalue of $\widetilde{\mathbf{X}}_2\widetilde{\mathbf{X}}_1^{-1}$ associated

with eigenvectors I_j and $z = \text{card}(I_j)$. The matrix $\mathbf{K}_j^{(n)}$ contains the z eigenvectors associated with the eigenvalues of $\mathbf{Y}_2^{(n)}(\mathbf{Y}_1^{(n)})^{-1}$ on the diagonal of $\mathbf{\Lambda}_j^{(n)}$. As $\underline{\mathbf{Y}}^{(n)} \to \widetilde{\underline{\mathbf{X}}}$, the eigenvectors $\mathbf{K}_j^{(n)}$ converge to linear dependence but they remain linearly independent of all other eigenvectors. This, together with $\|\mathbf{Y}_1^{(n)}\|$ being bounded, yields that $\|\mathbf{Q}_j^{(n)}\|$ in (3.5) is bounded. Since

$$\left\| \sum_{r \in I_j} \omega_r^{(n)} \left(\mathbf{a}_r^{(n)} \circ \mathbf{b}_r^{(n)} \circ \mathbf{c}_r^{(n)} \right) \right\| = \sqrt{\|\mathbf{P}_j^{(n)}\|^2 + \|\mathbf{Q}_j^{(n)}\|^2}\,, \tag{3.6}$$

it follows that the left-hand side of (3.6) is bounded. Hence, (1.10) holds and the sequence of rank-p decompositions of $\underline{\mathbf{Y}}^{(n)}$ will exhibit diverging CP components as $\underline{\mathbf{Y}}^{(n)} \to \widetilde{\underline{\mathbf{X}}}$. Moreover, the groups $I_1, \ldots, I_m$ of diverging CP components and the number of components in each group are related to the eigenvalues and eigenvectors of $\widetilde{\mathbf{X}}_2\widetilde{\mathbf{X}}_1^{-1}$ as we have seen above.

Case 8. We have $p > q$ and $R = q$. We assume that the optimal solution $\widetilde{\underline{\mathbf{X}}}$ of problem (2.2) lies in $\mathcal{U}_R^{(12)}$, which is the set $\mathcal{Q}_{R+1}$ in this case. The set $\mathcal{U}_R^{(12)}$ is defined by (2.18). This implies that for any nonsingular $\mathbf{S}$ such that

$$\mathbf{S}\,\widetilde{\mathbf{X}}_1 = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{O} \end{bmatrix} \qquad \text{and} \qquad \mathbf{S}\,\widetilde{\mathbf{X}}_2 = \begin{bmatrix} \mathbf{H}_2 \\ \mathbf{O} \end{bmatrix}, \tag{3.7}$$

the $R \times R$ matrix $\mathbf{H}_2\mathbf{H}_1^{-1}$ has R real eigenvalues (which are not all distinct) and is not diagonalizable. By Proposition 1.2, $\text{rank}_\circ(\widetilde{\underline{\mathbf{X}}}) = \text{rank}_\circ(\underline{\mathbf{H}}) \geq R + 1$. Let $\underline{\mathbf{Y}}^{(n)}$ be a sequence of CP updates converging to $\widetilde{\underline{\mathbf{X}}}$. For a fixed $\mathbf{S}$ in (3.7), let $\mathbf{S}^{(n)}$ be such that it is nonsingular for all n, $\mathbf{S}^{(n)} \to \mathbf{S}$ and

$$\mathbf{S}^{(n)}\,\mathbf{Y}_1^{(n)} = \begin{bmatrix} \mathbf{H}_1^{(n)} \\ \mathbf{O} \end{bmatrix} \qquad \text{and} \qquad \mathbf{S}^{(n)}\,\mathbf{Y}_2^{(n)} = \begin{bmatrix} \mathbf{H}_2^{(n)} \\ \mathbf{O} \end{bmatrix}. \tag{3.8}$$

Then $\text{rank}_\circ(\underline{\mathbf{H}}^{(n)}) \leq R$ and $\underline{\mathbf{H}}^{(n)} \to \underline{\mathbf{H}}$ as $\underline{\mathbf{Y}}^{(n)} \to \widetilde{\underline{\mathbf{X}}}$. Since $\underline{\mathbf{H}}^{(n)}$, $\underline{\mathbf{H}} \in \mathbb{R}^{R \times R \times 2}$, it follows from case 2 above that the sequence $\underline{\mathbf{H}}^{(n)}$ will exhibit diverging CP components as it converges to $\underline{\mathbf{H}}$. Denote the unique rank-R decomposition of $\underline{\mathbf{H}}^{(n)}$ by $(\mathbf{A}^{(n)}, \mathbf{B}^{(n)}, \mathbf{C}^{(n)}, \mathbf{\Omega}^{(n)})$. Then the rank-$R$ decomposition of $\underline{\mathbf{Y}}^{(n)}$ is

$$\left((\mathbf{S}^{(n)})^{-1} \begin{bmatrix} \mathbf{A}^{(n)} \\ \mathbf{O} \end{bmatrix}, \mathbf{B}^{(n)}, \mathbf{C}^{(n)}, \mathbf{\Omega}^{(n)} \right). \tag{3.9}$$

The k-rank of $\mathbf{A}^{(n)}$ equals the k-rank of the first component matrix in (3.9). Hence, by virtue of Kruskal's condition (1.7) also the rank-R decomposition (3.9) is unique. Moreover, the decomposition will exhibit the same pattern of diverging CP components as $(\mathbf{A}^{(n)}, \mathbf{B}^{(n)}, \mathbf{C}^{(n)}, \mathbf{\Omega}^{(n)})$ when $\underline{\mathbf{Y}}^{(n)} \to \widetilde{\underline{\mathbf{X}}}$. Note that $\text{rank}_\circ(\underline{\mathbf{H}}) = \text{rank}_\circ(\widetilde{\underline{\mathbf{X}}})$ and Proposition 1.2 imply that the number of groups of diverging CP components does not depend on $\mathbf{S}$.

Cases 3, 5, and 9. We have $p \geq q$ and $R < q$. We assume that the optimal solution $\widetilde{\underline{\mathbf{X}}}$ of problem (2.2) satisfies

$$\mathbf{S}\,\widetilde{\mathbf{X}}_1\,\mathbf{T} = \begin{bmatrix} \mathbf{G}_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} \qquad \text{and} \qquad \mathbf{S}\,\widetilde{\mathbf{X}}_2\,\mathbf{T} = \begin{bmatrix} \mathbf{G}_2 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, \tag{3.10}$$

where $\mathbf{S}$ and $\mathbf{T}$ are nonsingular and the $R \times R$ matrix $\mathbf{G}_2\mathbf{G}_1^{-1}$ has R real eigenvalues (which are not all distinct) and is not diagonalizable. By Proposition 1.2, $\text{rank}_\circ(\underline{\widetilde{\mathbf{X}}}) = \text{rank}_\circ(\underline{\mathbf{G}}) \geq R+1$. Let $\underline{\mathbf{Y}}^{(n)}$ be a sequence of CP updates converging to $\underline{\widetilde{\mathbf{X}}}$. For fixed $\mathbf{S}$ and $\mathbf{T}$ in (3.10), let $\mathbf{S}^{(n)}$ and $\mathbf{T}^{(n)}$ be such that they are nonsingular for all n, $\mathbf{S}^{(n)} \to \mathbf{S}$, $\mathbf{T}^{(n)} \to \mathbf{T}$, and

$$\mathbf{S}^{(n)}\,\mathbf{Y}_1^{(n)}\,\mathbf{T}^{(n)} = \begin{bmatrix} \mathbf{G}_1^{(n)} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} \qquad \text{and} \qquad \mathbf{S}^{(n)}\,\mathbf{Y}_2^{(n)}\,\mathbf{T}^{(n)} = \begin{bmatrix} \mathbf{G}_2^{(n)} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}. \tag{3.11}$$

Then $\text{rank}_\circ(\underline{\mathbf{G}}^{(n)}) \leq R$ and $\underline{\mathbf{G}}^{(n)} \to \underline{\mathbf{G}}$ as $\underline{\mathbf{Y}}^{(n)} \to \underline{\widetilde{\mathbf{X}}}$. Since $\underline{\mathbf{G}}^{(n)}$, $\underline{\mathbf{G}} \in \mathbb{R}^{R\times R\times 2}$, it follows from case 2 that the sequence $\underline{\mathbf{G}}^{(n)}$ will exhibit diverging CP components as it converges to $\underline{\mathbf{G}}$. Denote the unique rank-R decomposition of $\underline{\mathbf{G}}^{(n)}$ by $(\mathbf{A}^{(n)}, \mathbf{B}^{(n)}, \mathbf{C}^{(n)}, \mathbf{\Omega}^{(n)})$. Then the rank-$R$ decomposition of $\underline{\mathbf{Y}}^{(n)}$ is

$$\left((\mathbf{S}^{(n)})^{-1} \begin{bmatrix} \mathbf{A}^{(n)} \\ \mathbf{O} \end{bmatrix}, (\mathbf{T}^{(n)})^{-T} \begin{bmatrix} \mathbf{B}^{(n)} \\ \mathbf{O} \end{bmatrix}, \mathbf{C}^{(n)}, \mathbf{\Omega}^{(n)} \right). \tag{3.12}$$

The k-ranks of $\mathbf{A}^{(n)}$ and $\mathbf{B}^{(n)}$ equal the k-ranks of the first two component matrices in (3.12). Hence, by virtue of Kruskal's condition (1.7) the rank-R decomposition (3.12) is also unique. Moreover, the decomposition will exhibit the same pattern of diverging CP components as $(\mathbf{A}^{(n)}, \mathbf{B}^{(n)}, \mathbf{C}^{(n)}, \mathbf{\Omega}^{(n)})$ when $\underline{\mathbf{Y}}^{(n)} \to \underline{\widetilde{\mathbf{X}}}$. Note that $\text{rank}_\circ(\underline{\mathbf{G}}) = \text{rank}_\circ(\underline{\widetilde{\mathbf{X}}})$ and Proposition 1.2 imply that the number of groups of diverging CP components does not depend on $\mathbf{S}$ and $\mathbf{T}$.

4. Simulation results. Here, we illustrate the cases in Table 3.1 by trying to calculate (using a CP algorithm) a best rank-R approximation of random $p \times q \times 2$ arrays, the elements of which are sampled independently from the uniform distribution on $[-1, 1]$. We consider cases 3, 5, 8, and 9, in which we conjecture diverging CP components to occur on a set of positive volume, and case 7, in which we conjecture diverging CP components to occur on a set of zero volume. Simulation results in Stegeman [34] show that for case 2 diverging CP components always occur, which is in agreement with our conjecture in this case. Although different sampling distributions will give different results on the percentages of cases of diverging CP components, we feel that the outcomes presented below are useful to show that diverging CP components are a serious problem indeed. As a CP algorithm, we use the multilinear engine by Paatero [29].

The simulation results in Stegeman [34] have indicated that, for random $p \times p \times 2$ arrays $\underline{\mathbf{X}}$, problem (2.2) has a unique optimal solution $\underline{\widetilde{\mathbf{X}}}$. If $\underline{\widetilde{\mathbf{X}}} \in \mathcal{S}_R$, then problem (2.1) has a unique optimal solution and diverging CP components do not occur. If $\underline{\widetilde{\mathbf{X}}} \notin \mathcal{S}_R$, then it is approximated arbitrarily close by arrays in $\mathcal{S}_R$. Notice that if the component matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ have full k-rank (which they usually have), then Kruskal's condition (1.7) holds in cases 3, 5, 7, 8, and 9. Hence, if $\underline{\widetilde{\mathbf{X}}} \notin \mathcal{S}_R$, then the arrays in $\mathcal{S}_R$ close to $\underline{\widetilde{\mathbf{X}}}$ have a unique CP decomposition exhibiting diverging CP components.

The reasoning above implies that it suffices to use only one run of the CP algorithm for each array $\underline{\mathbf{X}}$ (with random starting values for the component matrices). For cases 3 and 5, we consider 100 random $4 \times 4 \times 2$ arrays $\underline{\mathbf{X}}$ and $R = 3$. The rank of $\underline{\mathbf{X}}$ depends on whether $\mathbf{X}_2\mathbf{X}_1^{-1}$ has 4 real eigenvalues (rank 4) or some complex eigenvalues (rank 5). In cases 8 and 9, we consider 100 random $5 \times 3 \times 2$ arrays and 50 random $5 \times 4 \times 2$ arrays, respectively, both with $R = 3$. For case 7, we consider

100 random $6 \times 3 \times 2$ arrays and $R = 5$. Table 4.1 gives the relative frequency of the occurrence of diverging CP components we encountered in these cases. As can be seen, diverging CP components occur quite often in cases 3, 5, 8, and 9, and they do not occur at all in case 7. Hence, the results in Table 4.1 are in agreement with our conjectures in Tabel 3.1 regarding the occurrence of diverging CP components.

TABLE 4.1
Results of calculating rank-R approximations to random $p \times q \times 2$ arrays in cases 3, 5, 7, 8, *and* 9 *of Table* 3.1.

Case	$\underline{\mathbf{X}} : p \times q \times 2$	rank($\underline{\mathbf{X}}$)	R	Diverging CP components
3	$4 \times 4 \times 2$	5	3	44 out of 70 (63%)
5	$4 \times 4 \times 2$	4	3	6 out of 30 (20%)
7	$6 \times 3 \times 2$	6	5	0 out of 100
8	$5 \times 3 \times 2$	5	3	51 out of 100 (51%)
9	$5 \times 4 \times 2$	5	3	24 out of 50 (48%)

5. Discussion. We have considered low-rank approximations to generic $p \times q \times 2$ arrays. For all combinations of p, q, and R, we presented conjectures on whether a best rank-R approximation exists almost everywhere, on a set of positive volume or on a set of zero volume. In the cases where no best rank-R approximation exists, this is due to the fact that the optimal boundary points of $\mathcal{S}_R$ (i.e., the optimal solutions of problem (2.2)) do not lie in $\mathcal{S}_R$ itself. We showed (under some regularity condition) that if a sequence of CP updates converges to such an optimal boundary point, it necessarily exhibits diverging CP components.

This explanation of diverging CP components confirms the statement of Kruskal, Harshman, and Lundy [24] that these occur due to the fact that the CP objective function does not attain its infimum, and that any sequence of CP updates of which the objective value is approaching the infimum must fail to converge and exhibits diverging CP components. Also, the concept of a sequence of CP updates converging to a boundary point $\widetilde{\underline{\mathbf{X}}} \notin \mathcal{S}_R$ can be found in Kruskal, Harshman, and Lundy [24] for the case $p = q = R = 2$. Whether diverging CP components always occur if there is no best rank-R approximation is still an open problem.

As in Stegeman [34], the occurrence of diverging CP components in the cases in Table 3.1 and their explanations are still valid when the Frobenius norm in the CP objective function is replaced by any other norm (e.g., weighted least squares or Gaussian maximum likelihood). This is because all norms on the finite-dimensional vector space are equivalent and induce the same (i.e., the Euclidian) topology.

Note that, as in Stegeman [34], the occurrence of diverging CP components in Table 3.1 does not depend on the algorithm used to minimize the CP objective function. Hence, modified CP algorithms designed to avoid diverging CP components are of no use here.

The diverging CP components in case 2 of Table 3.1 occur due to the two-valued typical rank of real-valued $p \times p \times 2$ arrays and the uniqueness of their rank-p decomposition; see the discussion in Stegeman [34]. Our results on diverging CP components for $p \times q \times 2$ arrays are based upon this. The typical rank of $p \times p \times 2$ arrays over the complex field is p. Therefore, the cases of diverging CP components described in this paper do not occur in the complex-valued CP model. However, also for the complex

field, a best low-rank approximation does not always exist. See the example in De Silva and Lim [10, Proposition 4.6], which carries over to the complex field.

In cases 3, 5, 8, and 9 of Table 3.1, diverging CP components occur due to the fact that we may transform the CP problem for $p \times q \times 2$ arrays to the lower-dimensional CP problem for $R \times R \times 2$ arrays, for which case 2 applies. This shows that a two-valued typical rank of the target array $\underline{\mathbf{X}}$ is not necessary for diverging CP components to occur. However, the two-valued typical rank for $R \times R \times 2$ arrays is still necessary for diverging CP components to occur.

Zijlstra and Kiers [43] observed that cases of two diverging CP components occur not only in CP but also in other variants of factor analysis. They show that two-way and three-way factor analysis models which yield diverging components necessarily have rotationally unique components. For the cases we have examined, diverging CP components always occur together with uniqueness of the CP solution. This raises the question whether (partial) uniqueness of the CP solution is necessary for diverging components to occur. Stegeman [35] has shown that this is not the case. Indeed, in the cases of $3 \times 3 \times 5$ arrays with symmetric slices and $8 \times 4 \times 3$ arrays, diverging CP components occur on a set of positive volume while the CP decompositions of the CP updates are not unique.

Acknowledgments. The author would like to thank the anonymous reviewers for valuable suggestions to improve the paper and for bringing to the author's attention the algebraic complexity literature dealing with array rank and border rank.

REFERENCES

[1] D. Bini, M. Capovani, F. Romani, and G. Lotti, *$O(n^{2.7799})$ complexity for $n \times n$ approximate matrix multiplication*, Inform. Process. Lett., 8 (1979), pp. 234–235.

[2] D. Bini, G. Lotti, and F. Romani, *Approximate solutions for the bilinear form computational problem*, SIAM J. Comput., 9 (1980), pp. 692–697.

[3] D. Bini, *Relations between exact and approximate bilinear algorithms. Applications*, Calcolo, 17 (1980), pp. 87–97.

[4] D. Bini, *Border rank of a $p \times q \times 2$ tensor and the optimal approximation of a pair of bilinear forms*, in Automata, Languages and Programming, J. W. de Bakker and J. van Leeuwen, eds., Lecture Notes in Comput. Sci. 85, Springer, New York, 1980, pp. 98–108.

[5] D. Bini, *Border rank of $m \times n \times (mn - q)$ tensors*, Linear Algebra Appl., 79 (1986), pp. 45–51.

[6] R. W. Brockett and D. Dobkin, *On the optimal evaluation of a set of bilinear forms*, Linear Algebra Appl., 19 (1978), pp. 207–235.

[7] P. Bürgisser, M. Clausen, and M. A. Shokrollahi, *Algebraic Complexity Theory*, Springer, Berlin, 1997.

[8] Y. Z. Cao, Z. P. Chen, C. Y. Mo, H. L. Wu, and R. Q. Yu, *A Parafac algorithm using penalty diagonalization error (PDE) for three-way data array resolution*, The Analyst, 125 (2000), pp. 2303–2310.

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

[10] V. De Silva and L.-H. Lim, *Tensor Rank and the Ill-Posedness of the Best Low-Rank Approximation Problem*, SCCM Technical report, 06-06, Stanford University, Palo Alto, CA, 2006; available online at http://www.sccm.stanford.edu/wrap/pubtech.html.

[11] C. Eckart and G. Young, *The approximation of one matrix by another of lower rank*, Psychometrika, 1 (1936), pp. 211–218.

[12] F. R. Gantmacher, *The Theory of Matrices*, Volumes 1 and 2, Chelsea, New York, 1959.

[13] R. A. Harshman, *Foundations of the Parafac procedure: Models and conditions for an "explanatory" multimodal factor analysis*, UCLA Working Papers in Phonetics, 16 (1970), pp. 1–84.

[14] R. A. Harshman and M. E. Lundy, *Data preprocessing and the extended Parafac model*, in Research Methods for Multimode Data Analysis, H. G. Law, C. W. Snyder Jr., J. A. Hattie,

and R. P. McDonald, eds., Praeger, New York, 1984, pp. 216–284.
[15] F. L. Hitchcock, *The expression of a tensor or a polyadic as a sum of products*, J. Math. Phys., 6 (1927), pp. 164–189.
[16] F. L. Hitchcock, *Multiple invariants and generalized rank of a p-way matrix or tensor*, J. Math. Phys., 7 (1927), pp. 39–79.
[17] P. K. Hopke, P. Paatero, H. Jia, R. T. Ross, and R. A. Harshman, *Three-way (Parafac) factor analysis: Examination and comparison of alternative computational methods as applied to ill-conditioned data*, Chemom. Intel. Lab. Syst., 43 (1998), pp. 25–42.
[18] J. Ja'Ja', *Optimal evaluation of pairs of bilinear forms*, SIAM J. Comput., 8 (1979), pp. 443–462.
[19] J. Ja'Ja', *An addendum to Kronecker's theory of pencils*, SIAM J. Appl. Math., 37 (1979), pp. 700–712.
[20] L. Kronecker, *Algebraische reduction der schaaren bilinearer formen*, Sitzungber. Akad. Berlin (1890), pp. 763–776.
[21] P. M. Kroonenberg, *Three-Mode Principal Component Analysis*, DSWO, Leiden, The Netherlands, 1983.
[22] J. B. Kruskal, *Three-way arrays: Rank and uniqueness of trilinear decompositions, with applications to arithmetic complexity and statistics*, Linear Algebra Appl., 18 (1977), pp. 95–138.
[23] J. B. Kruskal, *Rank, decomposition, and uniqueness for* 3*-way and N-way arrays*, in Multiway Data Analysis, R. Coppi and S. Bolasco, eds., North-Holland, Amsterdam, 1989, pp. 7–18.
[24] J. B. Kruskal, R. A. Harshman, and M. E. Lundy, *How* 3*-MFA data can cause degenerate Parafac solutions, among other relationships*, in Multiway Data Analysis, R. Coppi and S. Bolasco, eds., North-Holland, Amsterdam, 1989, pp. 115–121 .
[25] J. M. Landsberg, *The border rank of the multiplication of* 2×2 *matrices in seven*, J. Amer. Math. Soc., 19 (2006), pp. 447–459.
[26] L.-H. Lim, *What's possible and what's impossible in tensor decompositions/approximation*, Talk at the Tensor Decomposition Workshop, Palo Alto, CA, 2004; available online at http://csmr.ca.sandia.gov/~tgkolda/tdw2004/.
[27] L.-H. Lim, *Optimal solutions to nonnegative Parafac/multilinear NMF always exist*, Talk at the Workshop on Tensor Decompositions and Applications, CIRM, Luminy, Marseille, France, 2005; available online at http://www.etis.ensea.fr/~wtda/.
[28] B. C. Mitchell and D. S. Burdick, *Slowly converging Parafac sequences: Swamps and two-factor degeneracies*, J. Chemometrics, 8 (1994), pp. 155–168.
[29] P. Paatero, *The Multilinear Engine—A table-driven least squares program for solving multilinear problems, including the n-way Parallel Factor Analysis model*, J. Comput. Graph. Statist., 8 (1999), pp. 854–888.
[30] P. Paatero, *Construction and analysis of degenerate Parafac models*, J. Chemometrics, 14 (2000), pp. 285–299.
[31] W. S. Rayens and B. C. Mitchell, *Two-factor degeneracies and a stabilization of Parafac*, Chemometrics Intel. Lab. Syst., 38 (1997), pp. 173–181.
[32] N. D. Sidiropoulos, *Low-rank decomposition of multi-way arrays: A signal processing perspective*, Invited Plenary Lecture at the IEEE Sensor Array and Multichannel (SAM) Signal Processing Workshop, Sitges, Barcelona, Spain, 2004. Available online at http://www.telecom.tuc.gr/~nikos/.
[33] A. Smilde, R. Bro, and P. Geladi, *Multi-way Analysis: Applications in the Chemical Sciences*, John Wiley & Sons, New York, 2004.
[34] A. Stegeman, *Degeneracy in Candecomp/Parafac explained for* $p \times p \times 2$ *arrays of rank* $p+1$ *or higher*, Psychometrika, 71 (2006), pp. 483–501.
[35] A. Stegeman, *Degeneracy in Candecomp/Parafac explained for several three-sliced arrays with a two-valued typical rank*, Psychometrika, (2008), in press.
[36] V. Strassen, *Vermeidung von divisionen*, J. Reine Angew. Math., 246 (1973), pp. 184–202.
[37] J. M. F ten Berge, H. A. L. Kiers, and J. De Leeuw, *Explicit Candecomp/Parafac solutions for a contrived* $2 \times 2 \times 2$ *array of rank three*, Psychometrika, 53 (1988), pp. 579–584.
[38] J. M. F. ten Berge, *Kruskal's polynomial for* $2 \times 2 \times 2$ *arrays and a generalization to* $2 \times n \times n$ *arrays*, Psychometrika, 56 (1991), pp. 631–636.
[39] J. M. F. ten Berge and H. A. L. Kiers, *Simplicity of core arrays in three-way principal component analysis and the typical rank of* $p \times q \times 2$ *arrays*, Linear Algebra Appl., 294 (1999), pp. 169–179.
[40] J. M. F. ten Berge, *Partial uniqueness in Candecomp/Parafac*, J. Chemometrics, 18 (2004), pp. 12–16.

[41] G. Tomasi and R. Bro, *A Comparison of algorithms for fitting the Parafac model*, Comput. Statist. Data Anal., 50 (2006), pp. 1700–1734.
[42] K. Weierstrass, *Zur theorie der bilinearen und quadratischen formen*, Monatsber. Akad. Wiss. Berlin, (1868), pp. 310–338.
[43] B. J. H. Zijlstra and H. A. L. Kiers, *Degenerate solutions obtained from several variants of factor analysis*, J. Chemometrics, 16 (2002), pp. 596–605.