Are diverging CP components always nearly proportional?

Alwin Stegeman[†] and Lieven De Lathauwer[‡]

October 10, 2011

Abstract

Fitting a Candecomp/Parafac (CP) decomposition (also known as Canonical Polyadic decomposition) to a multi-way array or higher-order tensor, is equivalent to finding a best low-rank approximation to the multi-way array or higher-order tensor, where the rank is defined as the outer-product rank. However, such a best low-rank approximation may not exist due to the fact that the set of multi-way arrays with rank at most R is not closed for $R \geq 2$. Nonexistence of a best low-rank approximation results in (groups of) diverging rank-1 components when an attempt is made to compute the approximation. In this note, we show that in a group of two or three diverging components, the components converge to proportionality almost everywhere. A partial proof of this result for larger groups of diverging components is also given. Also, we give examples of groups of three, four, and six non-proportional diverging components. These examples are shown to be exceptional cases.

Keywords: tensor decomposition, low-rank approximation, Candecomp, Parafac, diverging components.

AMS subject classifications: 15A18, 15A22, 15A69, 49M27, 62H25.

[†]A. Stegeman is with the Heijmans Institute for Psychological Research, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands, phone: ++31 50 363 6193, fax: ++31 50 363 6304, email: a.w.stegeman@rug.nl, URL: http://www.gmw.rug.nl/∼stegeman. Research is supported by the Dutch Organisation for Scientific Research (NWO), VIDI grant 452-08-001.

[‡]L. De Lathauwer is with the Group Science, Engineering and Technology of the Katholieke Universiteit Leuven Campus Kortrijk, E. Sabbelaan 53, 8500 Kortrijk, Belgium, tel.: +32-(0)56-32.60.62, fax: +32-(0)56-24.69.99, e-mail: Lieven.DeLathauwer@kuleuven-kortrijk.be. He is also with the Department of Electrical Engineering (ESAT), Research Division SCD, of the Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium, tel.: +32-(0)16-32.86.51, fax: +32-(0)16-32.19.70, e-mail: Lieven.DeLathauwer@esat.kuleuven.be, URL: http://homes.esat.kuleuven.be/∼delathau/home.html. Research supported by: (1) Research Council K.U.Leuven: GOA-MaNet, CoE EF/05/006 Optimization in Engineering (OPTEC), CIF1 and STRT1/08/023 (2) F.W.O.: (a) project G.0427.10N, (b) Research Communities ICCoS, ANMMM and MLDM, (3) the Belgian Federal Science Policy Office: IUAP P6/04 (DYSCO, "Dynamical systems, control and optimization", 2007–2011), (4) EU: ERNSI.

1 Introduction

This note is an addendum to Stegeman [7] where the following subject is studied. Let $\circ$ denote the outer-product, and define the rank (over the real field) of $\mathcal{Y} \in \mathbb{R}^{I\times J\times K}$ as

$$\operatorname{rank}(\mathcal{Y}) = \min\{R \mid \mathcal{Y} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r\}\,. \tag{1.1}$$

Let

$$S_R(I,J,K) = \{\mathcal{Y} \in \mathbb{R}^{I\times J\times K} \mid \operatorname{rank}(\mathcal{Y}) \leq R\}\,, \tag{1.2}$$

be the set of $I \times J \times K$ arrays with at most rank R, and let $\overline{S}_R(I,J,K)$ denote the closure of $S_R(I,J,K)$, i.e. the union of the set itself and its boundary points in $\mathbb{R}^{I\times J\times K}$.

Let $\mathcal{Z} \in \mathbb{R}^{I\times J\times K}$ and $||\cdot||$ denote the Frobenius norm on $\mathbb{R}^{I\times J\times K}$. Consider the following low-rank approximation problem.

$$\min\{||\mathcal{Z} - \mathcal{Y}|| \mid \mathcal{Y} \in S_R(I,J,K)\}\,. \tag{1.3}$$

The variables in this problem are actually the rank-1 terms in the rank-R decomposition of $\mathcal{Y}$:

$$\mathcal{Y} = \sum_{r=1}^{R} \omega_r \, (\mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r)\,, \tag{1.4}$$

where vectors $\mathbf{a}_r, \mathbf{b}_r, \mathbf{c}_r$ have norm 1, $r = 1, \ldots, R$. The decomposition (1.4) is known as Candecomp/Parafac (CP) or Canonical Polyadic decomposition. Let $\mathbf{A} = [\mathbf{a}_1|\ldots|\mathbf{a}_R]$, $\mathbf{B} = [\mathbf{b}_1|\ldots|\mathbf{b}_R]$ and $\mathbf{C} = [\mathbf{c}_1|\ldots|\mathbf{c}_R]$.

Assuming $\operatorname{rank}(\mathcal{Z}) > R$, an optimal solution of (1.3) will be a boundary point of the set $S_R(I,J,K)$. However, the set $S_R(I,J,K)$ is not closed for $R \geq 2$, and problem (1.3) may not have an optimal solution due to this fact; see De Silva and Lim [2]. This results in (groups of) diverging rank-1 terms (also known as diverging CP components) in (1.4) when an attempt is made to compute a best rank-R approximation, see Krijnen, Dijkstra and Stegeman [4]. In such a case, the solution array $\mathcal{Y}$ converges to a boundary point $\mathcal{X}$ of $S_R(I,J,K)$ with $\operatorname{rank}(\mathcal{X}) > R$. In practice this has the following consequences: while running a CP algorithm to solve (1.3), the decrease of $||\mathcal{Z} - \mathcal{Y}||$ becomes very slow, and some (groups of) columns of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ become nearly linearly dependent, while the corresponding weights ω_r become large in magnitude. However, the sum of the corresponding rank-1 terms remains small and contributes to a better CP fit. More formally, a group of diverging CP components corresponds to an index set $D \subseteq \{1, \ldots, R\}$ such that

$$|\omega_r^{(n)}| \to \infty\,, \qquad \text{for all } r \in D\,, \tag{1.5}$$

$$\text{while} \quad \|\sum_{r\in D} \omega_r^{(n)} \, (\mathbf{a}_r^{(n)} \circ \mathbf{b}_r^{(n)} \circ \mathbf{c}_r^{(n)})\| \qquad \text{is bounded}\,, \tag{1.6}$$

where the superscript (n) denotes the n-th CP update of the iterative CP algorithm. More than one group of diverging components may exist. In that case (1.5)-(1.6) hold for the corresponding disjoint index sets.

In practice, a group of diverging components as described above is almost always such that the corresponding columns of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ become nearly identical up to sign. That is, the diverging components are nearly proportional. In this note, we focus on the question whether this is always true or not. In Section 2, we show that in a group of two or three diverging components, the corresponding columns of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ converge to rank-1 almost everywhere. In Section 4, we give a partial proof of this result for larger groups of diverging components. In Sections 3, 5, and 6, we give examples of groups of three, four, and six non-proportional diverging components, respectively. These examples are shown to be exceptional cases.

In this note, all arrays, matrices, vectors, and scalars are real-valued.

2 Groups of two or three diverging components

We need the following notation. A matrix form of the CP decomposition (1.4) is

$$\mathbf{Y}_k = \mathbf{A}\,\mathbf{C}_k\,\mathbf{B}^T\,, \qquad k = 1, \ldots, K\,, \tag{2.1}$$

where $\mathbf{Y}_k$ is the k-th $I \times J$ frontal slice of $\mathcal{Y}$, and $\mathbf{C}_k$ is the diagonal matrix with row k of $\mathbf{C}$ as its diagonal. In (2.1), the weights ω_r are absorbed into $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$.

We use $\mathcal{Y} = (\mathbf{S}, \mathbf{T}, \mathbf{U}) \cdot \mathcal{G}$ to denote the multilinear matrix multiplication of an array $\mathcal{G} \in \mathbb{R}^{R \times P \times Q}$ with matrices $\mathbf{S}$ $(I \times R)$, $\mathbf{T}$ $(J \times P)$, and $\mathbf{U}$ $(K \times Q)$. The result of the multiplication is an $I \times J \times K$ array $\mathcal{Y}$ with entries

$$y_{ijk} = \sum_{r=1}^{R} \sum_{p=1}^{P} \sum_{q=1}^{Q} s_{ir}\, t_{jp}\, u_{kq}\, g_{rpq}\,, \tag{2.2}$$

where s_{ir}, t_{jp}, and u_{kq} are entries of $\mathbf{S}$, $\mathbf{T}$, and $\mathbf{U}$, respectively. We refer to multiplication $(\mathbf{I}_I, \mathbf{I}_J, \mathbf{U}) \cdot \mathcal{G}$ with $\mathbf{U}$ nonsingular as a *slicemix* of $\mathcal{G}$. We say that $\mathcal{G}$ has a *nonsingular slicemix* if $I = J$ and the array $(\mathbf{I}_I, \mathbf{I}_I, \mathbf{U}) \cdot \mathcal{G}$ has a nonsingular frontal $I \times I$ slice for some nonsingular $\mathbf{U}$.

For later use, we state the following lemma.

Lemma 2.1 *For $R \leq \min(I, J, K)$ and $\mathcal{Y} \in \overline{S}_R(I, J, K)$, it holds that $\mathcal{Y} = (\mathbf{S}, \mathbf{T}, \mathbf{U}) \cdot \mathcal{G}$ for some $\mathbf{S}$, $\mathbf{T}$, $\mathbf{U}$ column-wise orthonormal, and some $\mathcal{G} \in \overline{S}_R(R, R, R)$ with all frontal slices upper triangular. Moreover, $\mathcal{Y} \in S_R(I, J, K)$ if and only if $\mathcal{G} \in S_R(R, R, R)$.*

Proof. See Stegeman [7, Lemma 2.2 (b)]. □

Our main result is the following.

Theorem 2.2 *Let* $\mathcal{Y}^{(n)} = (\mathbf{A}^{(n)}, \mathbf{B}^{(n)}, \mathbf{C}^{(n)}, \mathbf{\Omega}^{(n)})$ *be an* $I \times J \times K$ CP *decomposition, i.e.*

$$\mathcal{Y}^{(n)} = \sum_{r=1}^{R} \omega_r^{(n)} \left(\mathbf{a}_r^{(n)} \circ \mathbf{b}_r^{(n)} \circ \mathbf{c}_r^{(n)}\right), \tag{2.3}$$

with $\mathbf{A}^{(n)}$, $\mathbf{B}^{(n)}$, *and* $\mathbf{C}^{(n)}$ *having the length-*1 *vectors* $\mathbf{a}_r^{(n)}$, $\mathbf{b}_r^{(n)}$, *and* $\mathbf{c}_r^{(n)}$ *as columns, respectively, and* $\mathbf{\Omega}^{(n)} = \operatorname{diag}(\omega_1^{(n)}, \ldots, \omega_R^{(n)})$. *Let* $\mathcal{Y}^{(n)} \to \mathcal{X}$ *with* $\operatorname{rank}(\mathcal{X}) > R$, *and such that* $|\omega_r^{(n)}| \to \infty$ *for each* r.

(*i*) *If* $R = 2$, *then* $\mathbf{A}^{(n)}$, $\mathbf{B}^{(n)}$, *and* $\mathbf{C}^{(n)}$ *converge to rank-*1 *matrices.*

(*ii*) *If* $R = 3$, *then* $\mathbf{A}^{(n)}$, $\mathbf{B}^{(n)}$, *and* $\mathbf{C}^{(n)}$ *converge to rank-*1 *matrices for almost all* $\mathcal{X}$.

Proof. Krijnen et al. [4] show that $\mathbf{A}^{(n)}$, $\mathbf{B}^{(n)}$, and $\mathbf{C}^{(n)}$ converge to rank-deficient matrices. For $R = 2$, this completes the proof. Next, let $R = 3$ and suppose $\min(I, J, K) \geq 3$. We have $\mathcal{X} \in \overline{S}_3(I, J, K)$ and $\mathcal{X}$ is a boundary point of $S_3(I, J, K)$. By Lemma 2.1, there exist $\mathbf{L}$, $\mathbf{M}$, $\mathbf{N}$ with orthonormal columns such that $\mathcal{X} = (\mathbf{L}, \mathbf{M}, \mathbf{N}) \cdot \mathcal{G}$, where the $3 \times 3 \times 3$ array $\mathcal{G}$ has all frontal slices upper triangular. We have $\operatorname{rank}(\mathcal{G}) = \operatorname{rank}(\mathcal{X}) > 3$ and the rank-3 CP sequence $(\mathbf{L}^T, \mathbf{M}^T, \mathbf{N}^T) \cdot \mathcal{Y}^{(n)} \to \mathcal{G}$. Slightly abusing notation, we write $\mathcal{Y}^{(n)} = (\mathbf{A}^{(n)}, \mathbf{B}^{(n)}, \mathbf{C}^{(n)}) \to \mathcal{G}$, where the weights $\omega_r^{(n)}$ have been absorbed in $\mathbf{A}^{(n)}, \mathbf{B}^{(n)}, \mathbf{C}^{(n)}$, which are now 3×3 matrices.

We assume that $\mathcal{G}$ has a nonsingular slicemix. This is true for almost all $\mathcal{X}$, i.e., the subset of boundary points $\mathcal{X}$ with rank larger than 3, for which $\mathcal{G}$ does not have a nonsingular slicemix, has lower dimensionality than the set of boundary points with rank larger than 3. In fact, if $\mathcal{G}$ does not have a nonsingular slicemix, then its upper triangular slices have a zero on their diagonals in the same position.

We apply a slicemix to $\mathcal{G}$ such that its first slice is nonsingular. Next, we premultiply the slices of $\mathcal{G}$ by the inverse of its first slice. Then $\mathcal{G}$ is of the form

$$\mathcal{G} = [\mathbf{G}_1 \,|\, \mathbf{G}_2 \,|\, \mathbf{G}_3] = \left[\begin{array}{ccc|ccc|ccc} 1 & 0 & 0 & a & d & f & \alpha & \delta & \nu \\ 0 & 1 & 0 & 0 & b & e & 0 & \beta & \epsilon \\ 0 & 0 & 1 & 0 & 0 & c & 0 & 0 & \gamma \end{array}\right]. \tag{2.4}$$

Since a matrix cannot be approximated arbitrarily well by a matrix of lower rank, it follows that the approximating rank-3 sequence $\mathcal{Y}^{(n)}$ has a nonsingular slicemix for n large enough. Moreover, by Lemma 2.1 we may assume without loss of generality that $\mathcal{Y}^{(n)}$ has the form (2.4). We denote the entries of $\mathcal{Y}^{(n)}$ with subscript n, i.e. $a_n, \ldots, f_n$ and $\alpha_n, \ldots, \nu_n$. Hence,

$$\mathcal{Y}^{(n)} = [\mathbf{Y}_1^{(n)} \,|\, \mathbf{Y}_2^{(n)} \,|\, \mathbf{Y}_3^{(n)}] = \left[\begin{array}{ccc|ccc|ccc} 1 & 0 & 0 & a_n & d_n & f_n & \alpha_n & \delta_n & \nu_n \\ 0 & 1 & 0 & 0 & b_n & e_n & 0 & \beta_n & \epsilon_n \\ 0 & 0 & 1 & 0 & 0 & c_n & 0 & 0 & \gamma_n \end{array}\right]. \tag{2.5}$$

Next, we consider the rank-3 decomposition $(\mathbf{A}^{(n)}, \mathbf{B}^{(n)}, \mathbf{C}^{(n)})$ of $\mathcal{Y}^{(n)}$, which can be written as $\mathbf{Y}_k^{(n)} = \mathbf{A}^{(n)}\,\mathbf{C}_k^{(n)}\,(\mathbf{B}^{(n)})^T$, where diagonal matrix $\mathbf{C}_k^{(n)}$ has row k of $\mathbf{C}^{(n)}$ as its diagonal, $k = 1, 2, 3$. Since $\mathbf{Y}_1^{(n)} = \mathbf{I}_3$, matrices $\mathbf{A}^{(n)}$ and $\mathbf{B}^{(n)}$ are nonsingular. Without loss of generality, we set $\mathbf{C}_1^{(n)} = \mathbf{I}_3$. Then $(\mathbf{A}^{(n)})^{-1} = (\mathbf{B}^{(n)})^T$ and $\mathbf{Y}_k^{(n)} = \mathbf{A}^{(n)}\,\mathbf{C}_k^{(n)}\,(\mathbf{A}^{(n)})^{-1}$ for $k = 2, 3$. Hence, slices $\mathbf{Y}_2^{(n)}$ and $\mathbf{Y}_3^{(n)}$ have the same eigenvectors. Moreover, their three eigenvectors are linearly independent, and their eigenvalues are on the diagonals of $\mathbf{C}_2^{(n)}$ and $\mathbf{C}_3^{(n)}$, respectively. Since $\mathbf{Y}_2^{(n)}$ and $\mathbf{Y}_3^{(n)}$ have eigenvalues a_n, b_n, c_n and $\alpha_n, \beta_n, \gamma_n$, respectively, we obtain

$$\mathbf{C}^{(n)} = \begin{bmatrix} 1 & 1 & 1 \\ a_n & b_n & c_n \\ \alpha_n & \beta_n & \gamma_n \end{bmatrix}. \tag{2.6}$$

From Krijnen et al. [4] we know that $\mathbf{A}^{(n)}$, $\mathbf{B}^{(n)}$, and $\mathbf{C}^{(n)}$ converge to matrices with ranks less than 3. The eigendecomposition $\mathbf{Y}_k^{(n)} = \mathbf{A}^{(n)}\,\mathbf{C}_k^{(n)}\,(\mathbf{A}^{(n)})^{-1}$ converges to frontal slice $\mathbf{G}_k$ of $\mathcal{G}$, $k = 2, 3$. Hence, the eigenvectors in $\mathbf{A}^{(n)}$ converge to those of $\mathbf{G}_k$, $k = 2, 3$. Suppose $\mathbf{A}^{(n)}$ has a rank-1 limit. Then $\mathbf{G}_k$ has only one eigenvector and three identical eigenvalues, $k = 2, 3$. Hence, $a = b = c$ and $\alpha = \beta = \gamma$. Suppose $\mathbf{A}^{(n)}$ has a rank-2 limit $[\mathbf{a}_1\ \mathbf{a}_2\ \mathbf{a}_3]$, with $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ eigenvectors associated with eigenvalues a, b, c of $\mathbf{G}_2$, and eigenvalues α, β, γ of $\mathbf{G}_3$, respectively. Without loss of generality, let $\mathbf{a}_1$ and $\mathbf{a}_2$ be linearly independent. If $\mathbf{a}_3$ is proportional to either $\mathbf{a}_1$ or $\mathbf{a}_2$, then $\mathbf{B}^{(n)} = (\mathbf{A}^{(n)})^{-T}$ has large numbers in only two columns. This violates the assumption of $|\omega_r^{(n)}| \to \infty$ for $r = 1, 2, 3$. Hence, $\mathbf{a}_3$ is in the linear span of $\{\mathbf{a}_1, \mathbf{a}_2\}$ and not proportional to $\mathbf{a}_1$ or to $\mathbf{a}_2$. For an eigenvalue λ of $\mathbf{G}_k$, we define the eigenspace

$$E_k(\lambda) = \{\mathbf{x} \in \mathbb{R}^3 :\ \mathbf{G}_k\mathbf{x} = \lambda\,\mathbf{x}\}, \qquad k = 2, 3\,. \tag{2.7}$$

It holds that $\lambda_1 \neq \lambda_2$ implies $E_k(\lambda_1) \cap E_k(\lambda_2) = \{\mathbf{0}\}$. If a, b, c are distinct, then $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ would be linearly independent, which is not the case. Without loss of generality, let $a = b$. Then $\mathbf{a}_3 \in E_2(a) \cap E_2(c)$, which is impossible if $a \neq c$. Hence, it follows that $a = b = c$. The proof of $\alpha = \beta = \gamma$ is analogous. This implies that $\mathbf{C}^{(n)}$ in (2.6) converges to a rank-1 matrix.

As $\mathcal{Y}^{(n)} \to \mathcal{G}$, we first assume that the eigenvalues a_n, b_n, c_n are distinct and the eigenvalues $\alpha_n, \beta_n, \gamma_n$ are distinct. It can be verified that the eigenvectors $\mathbf{A}^{(n)}$ of $\mathbf{Y}_2^{(n)}$ associated with eigenvalues a_n, b_n, c_n are, respectively,

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 \\ \frac{b_n - a_n}{d_n} \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 \\ \frac{e_n(c_n - a_n)}{d_n e_n + f_n(c_n - b_n)} \\ \frac{(c_n - a_n)(c_n - b_n)}{d_n e_n + f_n(c_n - b_n)} \end{pmatrix}. \tag{2.8}$$

As explained above, the eigenvectors of $\mathbf{Y}_3^{(n)}$ (in terms of $\alpha_n, \ldots, \nu_n$) must be identical to those of $\mathbf{Y}_2^{(n)}$. We assume $d \neq 0$, $e \neq 0$, $f \neq 0$, $\delta \neq 0$, $\epsilon \neq 0$, $\nu \neq 0$, which holds for almost all $\mathcal{X}$. Combined

with $a = b = c$ and $\alpha = \beta = \gamma$, it follows from (2.8) that $\mathbf{A}^{(n)}$ converges to a rank-1 matrix. For $\mathbf{A}^{(n)}$ as in (2.8), the columns of $\mathbf{B}^{(n)} = (\mathbf{A}^{(n)})^{-T}$ are

$$\begin{pmatrix} 1 \\ \frac{d_n}{(a_n-b_n)} \\ \frac{d_n e_n + f_n(a_n-b_n)}{(a_n-b_n)(a_n-c_n)} \end{pmatrix}, \quad \begin{pmatrix} 0 \\ \frac{d_n}{(b_n-a_n)} \\ \frac{d_n e_n}{(a_n-b_n)(c_n-b_n)} \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ \frac{d_n e_n + f_n(c_n-b_n)}{(c_n-a_n)(c_n-b_n)} \end{pmatrix}. \tag{2.9}$$

Hence, each column of $\mathbf{B}^{(n)}$ contains large numbers for large n. After normalizing the third entries of each column of $\mathbf{B}^{(n)}$ to 1, we obtain

$$\begin{pmatrix} \frac{(a_n-b_n)(a_n-c_n)}{d_n e_n + f_n(a_n-b_n)} \\ \frac{d_n(a_n-c_n)}{d_n e_n + f_n(a_n-b_n)} \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ \frac{(b_n-c_n)}{e_n} \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{2.10}$$

It follows that also $\mathbf{B}^{(n)}$ converges to a rank-1 matrix.

Above, we assumed distinct eigenvalues a_n, b_n, c_n and $\alpha_n, \beta_n, \gamma_n$. Next, we show that cases with identical eigenvalues can be left out of consideration. We only consider cases where some of a_n, b_n, c_n are identical. Cases where some of α, β, γ are identical can be treated analogously. If $a_n = b_n \neq c_n$ for n large enough, then we must have $d_n = 0$ to obtain three linearly independent eigenvectors of $\mathbf{Y}_2^{(n)}$. This is due to the upper triangular form of $\mathbf{Y}_2^{(n)}$ in (2.5). This implies that $d = 0$ in the limit, which does not hold for almost all $\mathcal{X}$.

The case $a_n \neq b_n = c_n$ can be dealt with analogously. Here, we must have $e_n = 0$ to obtain three linearly independent eigenvectors of $\mathbf{Y}_2^{(n)}$ in (2.5). This implies that $e = 0$ in the limit, which does not hold for almost all $\mathcal{X}$.

Next, suppose $a_n = c_n \neq b_n$ for n large enough. To obtain three linearly independent eigenvectors of $\mathbf{Y}_2^{(n)}$ in (2.5), we must have $d_n e_n + f_n(c_n - b_n) = 0$. Since $c_n - b_n \to c - b = 0$, this implies that $de = 0$ in the limit, which does not hold for almost all $\mathcal{X}$.

Finally, we consider the case $a_n = b_n = c_n$ for n large enough. To obtain three linearly independent eigenvectors of $\mathbf{Y}_2^{(n)}$ in (2.5), we must have $d_n = e_n = f_n = 0$. This implies that $d = e = f = 0$ in the limit, which does not hold for almost all $\mathcal{X}$. This completes the proof for $\min(I, J, K) \geq 3$.

Next, let $\min(I, J, K) = 2$. Without loss of generality, we assume $I \geq J \geq K$. If $I = J = K = 2$, then $\mathcal{X}$ is a $2 \times 2 \times 2$ array, which has maximal rank 3 [3]. A contradiction to $\text{rank}(\mathcal{X}) > 3$. If $I > J = K = 2$, then by [2, Theorem 5.2] there exists a column-wise orthonormal $\mathbf{L}$ $(I \times 3)$ such that $\mathcal{X} = (\mathbf{L}, \mathbf{I}_2, \mathbf{I}_2) \cdot \mathcal{G}$, with $\mathcal{G}$ a $3 \times 2 \times 2$ array and $\text{rank}(\mathcal{X}) = \text{rank}(\mathcal{G})$. Since the maximal rank of $3 \times 2 \times 2$ arrays is 3 [3], we again obtain a contradiction.

Finally, let $I \geq J > K = 2$. By [2, Theorem 5.2] there exist column-wise orthonormal $\mathbf{L}$ $(I \times 3)$ and $\mathbf{M}$ $(J \times 3)$ such that $\mathcal{X} = (\mathbf{L}, \mathbf{M}, \mathbf{I}_2) \cdot \mathcal{G}$, with $\mathcal{G}$ a $3 \times 3 \times 2$ array and $\text{rank}(\mathcal{X}) = \text{rank}(\mathcal{G}) > 3$. The remainder of the proof is analogous to the beginning of the proof for $\min(I, J, K) \geq 3$. We

let $\mathcal{Y}^{(n)} = (\mathbf{A}^{(n)}, \mathbf{B}^{(n)}, \mathbf{C}^{(n)}) \to \mathcal{G}$, where $\mathbf{A}^{(n)}$ and $\mathbf{B}^{(n)}$ are 3×3, and $\mathbf{C}^{(n)}$ is 2×3. As shown in [6], array $\mathcal{G}$ can be transformed to have upper triangular slices. Assuming $\mathcal{G}$ has a nonsingular slicemix, we transform it to

$$\mathcal{G} = [\mathbf{G}_1 \,|\, \mathbf{G}_2] = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & a & d & f \\ 0 & 1 & 0 & 0 & b & e \\ 0 & 0 & 1 & 0 & 0 & c \end{array}\right]. \tag{2.11}$$

We assume $\mathcal{Y}^{(n)}$ to be of the same form, with entries $a_n, b_n, c_n, d_n, e_n, f_n$. As above, we have $\mathbf{B}^{(n)} = (\mathbf{A}^{(n)})^{-T}$,

$$\mathbf{C}^{(n)} = \begin{bmatrix} 1 & 1 & 1 \\ a_n & b_n & c_n \end{bmatrix}, \tag{2.12}$$

and eigendecomposition $\mathbf{Y}_2^{(n)} = \mathbf{A}^{(n)}\, \mathbf{C}_2^{(n)}\, (\mathbf{A}^{(n)})^{-1}$ converging to frontal slice $\mathbf{G}_2$. Krijnen et al. [4] show that $\mathbf{C}^{(n)}$ converges to a rank-deficient matrix. Hence, $\mathbf{C}^{(n)}$ converges to a rank-1 matrix, and we have $a = b = c$ in the limit.

We assume distinct eigenvalues a_n, b_n, c_n. As above, having some identical eigenvalues for n large enough yields $d = 0$ or $e = 0$ or $f = 0$, which does not hold for almost all $\mathcal{X}$. The eigenvectors $\mathbf{A}^{(n)}$ of $\mathbf{Y}_2^{(n)}$ are given by (2.8). We assume $d \neq 0$, $e \neq 0$, $f \neq 0$. Since $a = b = c$, it can be seen that $\mathbf{A}^{(n)}$ converges to a rank-1 matrix. The same is true for $\mathbf{B}^{(n)} = (\mathbf{A}^{(n)})^{-T}$, which has columns equal to (2.10) after normalizing the third entry of each column to 1. This completes the proof. □

3 Example of non-proportional diverging components for $R = 3$

The example is of size $3 \times 3 \times 2$. Let

$$\mathcal{X} = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & a & 0 & f \\ 0 & 1 & 0 & 0 & a & e \\ 0 & 0 & 1 & 0 & 0 & a \end{array}\right], \tag{3.1}$$

with $e \neq 0$ and $f \neq 0$. Here, $\mathcal{X}$ plays the role of $\mathcal{G}$ in (2.11). Since $d = 0$ in $\mathcal{X}$ above, this is an exception to almost all boundary arrays $\mathcal{X}$ with $\operatorname{rank}(\mathcal{X}) > 3$. We have $\mathcal{Y}^{(n)} = (\mathbf{A}^{(n)}, \mathbf{B}^{(n)}, \mathbf{C}^{(n)}) \to \mathcal{X}$, with

$$\mathbf{A}^{(n)} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & \frac{e_n\,(c_n - a_n)}{f_n\,(c_n - b_n)} \\ 0 & 0 & \frac{(c_n - a_n)}{f_n} \end{bmatrix}, \qquad \mathbf{B}^{(n)} = (\mathbf{A}^{(n)})^{-T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{f_n}{(a_n - c_n)} & \frac{e_n}{(b_n - c_n)} & \frac{f_n}{(c_n - a_n)} \end{bmatrix}, \tag{3.2}$$

and $\mathbf{C}^{(n)}$ as in (2.12). Let $a_n = a + 1/n$, $b_n = a - 1/n$, $c_n = a + 2/n$. Then

$$\mathbf{A}^{(n)} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & \frac{e}{3f} \\ 0 & 0 & 0 \end{bmatrix}, \qquad \mathbf{B}^{(n)} \to \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \tag{3.3}$$

where the columns of $\mathbf{B}^{(n)}$ are normalized such that their third entries are equal to 1. As we see, $\mathbf{A}^{(n)}$ converges to a rank-2 matrix. Note that $|\omega_r^{(n)}| \to \infty$ for $r = 1, 2, 3$ in this example, since all three columns of $\mathbf{B}^{(n)}$ in (3.2) will have large numbers as entries.

4 Extension to groups of four or more diverging components

Can statement (*ii*) of Theorem 2.2 also be proven for $R \geq 4$ if the requirement $R \leq \min(I, J, K)$ is added? A proof of this could be analogous to the proof for $R = 3$ under this requirement. That is, we have $\mathcal{Y}^{(n)} = (\mathbf{A}^{(n)}, \mathbf{B}^{(n)}, \mathbf{C}^{(n)}) \to \mathcal{G}$, where $\mathcal{G}$ is $R \times R \times R$ and has its first slice equal to $\mathbf{I}_R$ and its other slices upper triangular. Matrices $\mathbf{A}^{(n)}$, $\mathbf{B}^{(n)}$, and $\mathbf{C}^{(n)}$ have size $R \times R$, with $\mathbf{B}^{(n)} = (\mathbf{A}^{(n)})^{-T}$, $\mathbf{C}_1^{(n)} = \mathbf{I}_R$, and $\mathbf{Y}_k^{(n)} = \mathbf{A}^{(n)}\,\mathbf{C}_k^{(n)}\,(\mathbf{A}^{(n)})^{-1}$ converging to frontal slice $\mathbf{G}_k$ of $\mathcal{G}$, $k = 2, \ldots, R$. Suppose we have shown that $\mathbf{G}_k$, $k = 2, \ldots, R$, all have R identical eigenvalues. Then the limit of $\mathbf{C}^{(n)}$ analogous to (2.6) is a rank-1 matrix. For any fixed $R \geq 4$, we can use symbolic computation software to obtain expressions for $\mathbf{A}^{(n)}$ and $\mathbf{B}^{(n)}$ analogous to (2.8)-(2.10). This would imply that also $\mathbf{A}^{(n)}$ and $\mathbf{B}^{(n)}$, when normalized to have length-1 columns, converge to rank-1 matrices. The cases where $\mathbf{Y}_k^{(n)}$ has some identical eigenvalues for n large enough, $2 \leq k \leq R$, restrict some entries of $\mathcal{G}$ to zero. This is an exception to almost all boundary arrays $\mathcal{X}$ with $\text{rank}(\mathcal{X}) > R$.

The difficulty in obtaining the proof sketched above seems to be in showing that $\mathbf{G}_k$, $k = 2, \ldots, R$, all have R identical eigenvalues. For this to be proven, the possibilities for the rank and dependence structure of the limit of $\mathbf{A}^{(n)}$ must be analyzed. To have a group of R diverging components with $R \geq 4$, it must not only be checked that all columns of $\mathbf{B}^{(n)} = (\mathbf{A}^{(n)})^{-T}$ contain large numbers, but also that the R components do not consist of several different groups of diverging components. For example, if $R = 4$, then having two groups of two diverging components corresponds to two times two identical eigenvalues in the limit. In this case, the limit of $\mathbf{A}^{(n)}$ has two groups of two proportional columns.

To demonstrate the above, consider the case $R = 4$. It suffices to show that $\mathbf{G}_2$ has four identical eigenvalues. Let $\mathbf{G}_2$ have eigenvalues $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ and associated eigenvectors $\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2\ \mathbf{a}_3\ \mathbf{a}_4]$, with $\mathbf{A}^{(n)} \to \mathbf{A}$. Let $E(\lambda) = \{\mathbf{x} \in \mathbb{R}^4 :\ \mathbf{G}_2\mathbf{x} = \lambda\,\mathbf{x}\}$ denote the eigenspace corresponding to eigenvalue λ. We have $\text{rank}(\mathbf{A}) < 4$. If $\text{rank}(\mathbf{A}) = 1$, then $\mathbf{G}_2$ has only one eigenvector and four identical eigenvalues: $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4$.

Next, let $\text{rank}(\mathbf{A}) = 2$. Without loss of generality, we assume $\mathbf{a}_3, \mathbf{a}_4 \in \text{span}\{\mathbf{a}_1, \mathbf{a}_2\}$, with $\mathbf{a}_1$ and $\mathbf{a}_2$ linearly independent. Suppose $\lambda_1 = \lambda_2$. Then $\mathbf{a}_4 \in E(\lambda_4) \cap E(\lambda_1)$, which implies $\lambda_1 = \lambda_4$. Analogously, $\mathbf{a}_3 \in E(\lambda_3) \cap E(\lambda_1)$ implies $\lambda_1 = \lambda_3$. Hence, we obtain $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4$. Next, suppose $\lambda_1 \neq \lambda_2$. Because $\text{rank}(\mathbf{A}) = 2$, we have at most two distinct eigenvalues. If $\lambda_3 = \lambda_1 \neq \lambda_2 = \lambda_4$, then $\mathbf{a}_1$ and $\mathbf{a}_3$ are proportional and $\mathbf{a}_2$ and $\mathbf{a}_4$ are proportional. Hence, this is a case of two groups of two diverging components, and not one group of four diverging

components. If $\lambda_1 \neq \lambda_2 = \lambda_3 = \lambda_4$, then $\mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$ are proportional, and we have a group of three diverging components only (i.e., large numbers in three columns of $\mathbf{B}^{(n)} = (\mathbf{A}^{(n)})^{-T}$ only). Other possibilities for $\lambda_1 \neq \lambda_2$ and $\text{rank}(\mathbf{A}) = 2$ are analogous. It follows that if $\text{rank}(\mathbf{A}) = 2$, then $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4$.

Next, let $\text{rank}(\mathbf{A}) = 3$. Without loss of generality, we assume $\mathbf{a}_4 \in \text{span}\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$, with $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ linearly independent. Suppose $\lambda_1 = \lambda_2 = \lambda_3$. Then $\mathbf{a}_4 \in E(\lambda_4) \cap E(\lambda_1)$, which implies $\lambda_1 = \lambda_4$, and yields the desired result. Next, suppose $\lambda_1 = \lambda_2 \neq \lambda_3$. If $\lambda_4 = \lambda_1$, then we have a group of three diverging components only. If $\lambda_4 = \lambda_3$, then $\mathbf{a}_3$ and $\mathbf{a}_4$ are proportional, and we have a group of two diverging components only. If $\lambda_4 \neq \lambda_1$ and $\lambda_4 \neq \lambda_3$, then $\text{rank}(\mathbf{A}) = 4$ which is not possible. Next, suppose that $\lambda_1, \lambda_2, \lambda_3$ are distinct. Then λ_4 must be equal to one of them. Let $\lambda_4 = \lambda_1$. Then $\mathbf{a}_1$ and $\mathbf{a}_4$ are proportional, and we have a group of two diverging components only. Other possibilities for the equality of some eigenvalues can be treated analogously. It follows that if $\text{rank}(\mathbf{A}) = 3$, then $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4$.

As a final remark, we state that under the requirement $R \leq \min(I, J)$, the case $K = 2$ can be proven for any $R \geq 4$ analogous to the proof of $I \geq J > K = 2$ and $R = 3$.

5 Example of non-proportional diverging components for $R = 4$

This example is in the spirit of [2, Section 4]. Let $\min(I, J) \geq 6$, $K \geq 5$, and $R = 4$. Let $\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2]$ $(I \times 2)$ and $\mathbf{B} = [\mathbf{b}_1\ \mathbf{b}_2]$ $(J \times 2)$ be random matrices. For random $\mathbf{Q}$ (2×2), let $\tilde{\mathbf{A}} = [\tilde{\mathbf{a}}_1\ \tilde{\mathbf{a}}_2] = \mathbf{AQ}$ and $\tilde{\mathbf{B}} = [\tilde{\mathbf{b}}_1\ \tilde{\mathbf{b}}_2] = \mathbf{BQ}^{-T}$. Then $\mathbf{AB}^T = \tilde{\mathbf{A}}\tilde{\mathbf{B}}^T$, which implies

$$\mathbf{a}_1 \circ \mathbf{b}_1 \circ \mathbf{c} + \mathbf{a}_2 \circ \mathbf{b}_2 \circ \mathbf{c} - \tilde{\mathbf{a}}_1 \circ \tilde{\mathbf{b}}_1 \circ \mathbf{c} - \tilde{\mathbf{a}}_2 \circ \tilde{\mathbf{b}}_2 \circ \mathbf{c} = \mathcal{O}\,, \tag{5.1}$$

for any vector $\mathbf{c}$, where $\mathcal{O}$ denotes an allzero array. Let $\mathbf{X} = [\mathbf{x}_1\ \ldots\ \mathbf{x}_4]$ $(I \times 4)$ and $\mathbf{Y} = [\mathbf{y}_1\ \ldots\ \mathbf{y}_4]$ $(J \times 4)$ and $\mathbf{Z} = [\mathbf{z}_1\ \ldots\ \mathbf{z}_4]$ $(K \times 4)$ be random matrices.

We define

$$\mathbf{A}^{(n)} = \left[\, \mathbf{a}_1 + \frac{1}{n}\mathbf{x}_1 \,\middle|\, \mathbf{a}_2 + \frac{1}{n}\mathbf{x}_2 \,\middle|\, -\tilde{\mathbf{a}}_1 - \frac{1}{n}\mathbf{x}_3 \,\middle|\, -\tilde{\mathbf{a}}_2 - \frac{1}{n}\mathbf{x}_4 \,\right], \tag{5.2}$$

$$\mathbf{B}^{(n)} = \left[\, \mathbf{b}_1 + \frac{1}{n}\mathbf{y}_1 \,\middle|\, \mathbf{b}_2 + \frac{1}{n}\mathbf{y}_2 \,\middle|\, \tilde{\mathbf{b}}_1 + \frac{1}{n}\mathbf{y}_3 \,\middle|\, \tilde{\mathbf{b}}_2 + \frac{1}{n}\mathbf{y}_4 \,\right], \tag{5.3}$$

$$\mathbf{C}^{(n)} = \left[\, \mathbf{c} + \frac{1}{n}\mathbf{z}_1 \,\middle|\, \mathbf{c} + \frac{1}{n}\mathbf{z}_2 \,\middle|\, \mathbf{c} + \frac{1}{n}\mathbf{z}_3 \,\middle|\, \mathbf{c} + \frac{1}{n}\mathbf{z}_4 \,\right]. \tag{5.4}$$

If we let $\mathcal{Y}^{(n)} = n\,(\mathbf{A}^{(n)}, \mathbf{B}^{(n)}, \mathbf{C}^{(n)}) \to \mathcal{X}$, then by using (5.1) we obtain

$$\begin{aligned} \mathcal{X} \;=\; & \mathbf{a}_1 \circ \mathbf{b}_1 \circ \mathbf{z}_1 + \mathbf{a}_1 \circ \mathbf{y}_1 \circ \mathbf{c} + \mathbf{x}_1 \circ \mathbf{b}_1 \circ \mathbf{c} + \mathbf{a}_2 \circ \mathbf{b}_2 \circ \mathbf{z}_2 + \mathbf{a}_2 \circ \mathbf{y}_2 \circ \mathbf{c} + \mathbf{x}_2 \circ \mathbf{b}_2 \circ \mathbf{c} \\ & - \tilde{\mathbf{a}}_1 \circ \tilde{\mathbf{b}}_1 \circ \mathbf{z}_3 - \tilde{\mathbf{a}}_1 \circ \mathbf{y}_3 \circ \mathbf{c} - \mathbf{x}_3 \circ \tilde{\mathbf{b}}_1 \circ \mathbf{c} - \tilde{\mathbf{a}}_2 \circ \tilde{\mathbf{b}}_2 \circ \mathbf{z}_4 - \tilde{\mathbf{a}}_2 \circ \mathbf{y}_4 \circ \mathbf{c} - \mathbf{x}_4 \circ \tilde{\mathbf{b}}_2 \circ \mathbf{c}\,. \end{aligned} \tag{5.5}$$

The mode-1 rank of $\mathcal{X}$ equals $\text{rank}[\mathbf{A} \,|\, \tilde{\mathbf{A}} \,|\, \mathbf{X}] = 6$. The mode-2 rank of $\mathcal{X}$ equals $\text{rank}[\mathbf{B} \,|\, \tilde{\mathbf{B}} \,|\, \mathbf{Y}] = 6$. The mode-3 rank of $\mathcal{X}$ equals $\text{rank}[\mathbf{c} \,|\, \mathbf{Z}] = 5$. Since $\text{rank}(\mathcal{X})$ is at least equal to its mode-i rank,

$i = 1, 2, 3$, it follows that $\text{rank}(\mathcal{X}) \geq 6$. As we see, both $\mathbf{A}^{(n)}$ and $\mathbf{B}^{(n)}$ converge to rank-2 matrices, while $\mathbf{C}^{(n)}$ converges to a rank-1 matrix.

We create $\mathcal{X}$ according to (5.5) and compute the $4 \times 4 \times 4$ array $\mathcal{G}$ with upper triangular slices in the transformation $\mathcal{X} = (\mathbf{L}, \mathbf{M}, \mathbf{N}) \cdot \mathcal{G}$, using the Jacobi-type SGSD algorithm of [1] modified as described in [5]. Recall that existence of this transformation follows from Lemma 2.1. Next, we transform the slices of $\mathcal{G}$ such that its first slice becomes $\mathbf{I}_4$. We observe the following form for slices 2,3,4 of the transformed $\mathcal{G}$:

$$\begin{bmatrix} a & 0 & e & g \\ 0 & a & c & f \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & a \end{bmatrix}. \tag{5.6}$$

Hence, the slices have four identical eigenvalues and zeros in positions (1,2) and (3,4). The latter property shows that the limit point is an exception to almost all boundary arrays $\mathcal{X}$ with $\text{rank}(\mathcal{X}) > 4$.

6 Example of non-proportional diverging components for $R = 6$

This example is similar to the one in Section 5. Let $\min(I, J, K) \geq 8$ and $R = 6$. Let

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \tag{6.1}$$

$$\tilde{\mathbf{A}} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & -1 \end{bmatrix}, \qquad \tilde{\mathbf{B}} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad \tilde{\mathbf{C}} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \tag{6.2}$$

Then we have

$$\sum_{r=1}^{3} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r = \sum_{r=1}^{3} \tilde{\mathbf{a}}_r \circ \tilde{\mathbf{b}}_r \circ \tilde{\mathbf{c}}_r = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right], \tag{6.3}$$

where the latter is a rank-2 array. Let $\mathbf{S}$ $(I \times 2)$, $\mathbf{T}$ $(J \times 2)$, and $\mathbf{U}$ $(K \times 2)$ be random matrices. Let $\mathbf{X}$ $(I \times 6)$, $\mathbf{Y}$ $(J \times 6)$, and $\mathbf{Z}$ $(K \times 6)$ be random matrices. We define

$$\mathbf{A}^{(n)} = [\mathbf{S}\,\mathbf{A} \mid -\mathbf{S}\,\tilde{\mathbf{A}}] + \frac{1}{n}\,\mathbf{X} \cdot \text{diag}(1, 1, 1, -1, -1, -1)\,, \tag{6.4}$$

$$\mathbf{B}^{(n)} = [\mathbf{T}\,\mathbf{B} \mid \mathbf{T}\,\tilde{\mathbf{B}}] + \frac{1}{n}\,\mathbf{Y}\,, \qquad \mathbf{C}^{(n)} = [\mathbf{U}\,\mathbf{C} \mid \mathbf{U}\,\tilde{\mathbf{C}}] + \frac{1}{n}\,\mathbf{Z}\,. \tag{6.5}$$

Since the CP decompositions $(\mathbf{SA}, \mathbf{TB}, \mathbf{UC})$ and $(\mathbf{S}\tilde{\mathbf{A}}, \mathbf{T}\tilde{\mathbf{B}}, \mathbf{U}\tilde{\mathbf{C}})$ yield the same $I \times J \times K$ array, it follows that $\mathcal{Y}^{(n)} = n\,(\mathbf{A}^{(n)}, \mathbf{B}^{(n)}, \mathbf{C}^{(n)}) \to \mathcal{X}$, with

$$\begin{aligned} \mathcal{X} \;=\; & \sum_{r=1}^{3} (\mathbf{S}\mathbf{a}_r \circ \mathbf{T}\mathbf{b}_r \circ \mathbf{z}_r + \mathbf{S}\mathbf{a}_r \circ \mathbf{y}_r \circ \mathbf{U}\mathbf{c}_r + \mathbf{x}_r \circ \mathbf{T}\mathbf{b}_r \circ \mathbf{U}\mathbf{c}_r) \\ & - \sum_{s=1}^{3} (\mathbf{S}\tilde{\mathbf{a}}_s \circ \mathbf{T}\tilde{\mathbf{b}}_s \circ \mathbf{z}_{s+3} + \mathbf{S}\tilde{\mathbf{a}}_s \circ \mathbf{y}_{s+3} \circ \mathbf{U}\tilde{\mathbf{c}}_s + \mathbf{x}_{s+3} \circ \mathbf{T}\tilde{\mathbf{b}}_s \circ \mathbf{U}\tilde{\mathbf{c}}_s)\,. \end{aligned} \tag{6.6}$$

The mode-1 rank of $\mathcal{X}$ equals rank$[\mathbf{SA} \,|\, \mathbf{S\tilde{A}} \,|\, \mathbf{X}] = 8$. The mode-2 rank of $\mathcal{X}$ equals rank$[\mathbf{TB} \,|\, \mathbf{T\tilde{B}} \,|\, \mathbf{Y}] = 8$. The mode-3 rank of $\mathcal{X}$ equals rank$[\mathbf{UC} \,|\, \mathbf{U\tilde{C}} \,|\, \mathbf{Z}] = 8$. Since rank$(\mathcal{X})$ is at least equal to its mode-i rank, $i = 1, 2, 3$, it follows that rank$(\mathcal{X}) \geq 8$. As we see, matrices $\mathbf{A}^{(n)}$, $\mathbf{B}^{(n)}$, and $\mathbf{C}^{(n)}$ all converge to rank-2 matrices.

We create $\mathcal{X}$ according to (6.6) and compute the $6 \times 6 \times 6$ array $\mathcal{G}$ with upper triangular slices in the transformation $\mathcal{X} = (\mathbf{L}, \mathbf{M}, \mathbf{N}) \cdot \mathcal{G}$, using the Jacobi-type SGSD algorithm of [1] modified as described in [5]. The resulting slices of $\mathcal{G}$ have entries (4,4) and (5,5) almost zero. Hence, $\mathcal{G}$ does not seem to have a nonsingular slicemix. This implies that the limit point is an exception to almost all boundary arrays $\mathcal{X}$ with rank$(\mathcal{X}) > 6$.

References

[1] De Lathauwer, L, De Moor, B., & Vandewalle, J. (2004). Computation of the canconical decomposition by means of a simultaneous generalized Schur decomposition. *SIAM Journal on Matrix Analysis and Applications*, **26**, 295–327.

[2] De Silva, V., & Lim, L.-H. (2008). Tensor rank and the ill-posedness of the best low-rank approximation problem. *SIAM Journal on Matrix Analysis and Applications*, **30**, 1084–1127.

[3] Ja' Ja', J. (1979). Optimal evaluation of pairs of bilinear forms. *SIAM Journal on Computing*, **8**, 443–462.

[4] Krijnen, W.P., Dijkstra, T.K., & Stegeman, A. (2008). On the non-existence of optimal solutions and the occurrence of "degeneracy" in the Candecomp/Parafac model. *Psychometrika*, **73**, 431–439.

[5] Stegeman, A., & De Lathauwer, L. (2009). A method to avoid diverging components in the Candecomp/Parafac model for generic $I \times J \times 2$ arrays. *SIAM Journal on Matrix Analysis and Applications*, **30**, 1614–1638.

[6] Stegeman, A. (2010). The Generalized Schur Decomposition and the rank-R set of real $I \times J \times 2$ arrays. Technical Report, available online as `arXiv:1011.3432`

[7] Stegeman, A. (2011). Candecomp/Parafac - from diverging components to a decomposition in block terms. Technical Report, submitted.