CORE CONSISTENCY DIAGNOSTIC AIDED BY RECONSTRUCTION ERROR FOR ACCURATE ENUMERATION OF THE NUMBER OF COMPONENTS IN PARAFAC MODELS

Kefei Liu[1]*, H.C. So*[1]*, João Paulo C. L. da Costa*[2] *and Lei Huang*[3]

[1]Department of Electronic Engineering, City University of Hong Kong, Hong Kong
[2]Department of Electrical Engineering, University of Brasília, Brasília, Brazil
[3]Department of Electronic and Information Engineering,
Harbin Institute of Technology Shenzhen Graduate School, Shenzhen, China
E-mail: kefeilau@gmail.com, hcso@ee.cityu.edu.hk, jpdacosta@unb.br, lhuang@hitsz.edu.cn

ABSTRACT

Recently, the CORe CONsistency DIAgnostic (CORCONDIA) has attracted more and more attention as an effective tool for determining the number of components in parallel factor analysis (PARAFAC) or Tucker 3 models. In CORCONDIA, a proper user-defined threshold is required to ensure reliable performance. The optimal threshold increases with the signal-to-noise ratio (SNR), which results in significant probability of over-enumeration of the number of components for high SNRs under fixed threshold settings. We propose to first use a threshold interval to obtain lower and upper bounds of the estimates. The estimate takes the upper bound as its initial value and is then refined based on a sequence of hypothesis tests by exploiting the reconstruction error of the PARAFAC decomposition. The proposed scheme provides accurate detection for both low and high SNRs at almost no extra computational cost.

***Index Terms*—** Source enumeration, parallel factor analysis (PARAFAC), core consistency, multi-linear algebra

1. INTRODUCTION

The parallel factor analysis (PARAFAC) model [1, 2] has a variety of applications in chemometrics [3], blind source separation [4], multiple-input multiple-output (MIMO) radar [5] and wireless communications. In the PARAFAC model, a tensor is decomposed into the sum of rank-one tensors, which are defined as the outer product of vectors. Each rank-one factor corresponds to a signal, or component. Estimation of the tensor rank or number of signals or principal components in the underlying PARAFAC model of noisy R-D measurements, where $R \geq 3$, is an essential task.

Recently, the CORe CONsistency DIAgnostic (CORCONDIA) has been suggested for determining the number of components in PARAFAC [3] or Tucker 3 models [6]. For each candidate rank, the factor matrices of the PARAFAC model are first estimated via alternating least squares (ALS) PARAFAC decomposition, and then are used to calculate the core tensor. The core consistency defined as the variation of the estimated core relative to the ideal one, e.g., the identity tensor, is calculated. The tensor rank is determined by choosing the highest number of components that yields a core consistency value greater than a pre-defined threshold. Although being computationally expensive due to the iterative ALS algorithm, the CORCONDIA is an effective tool for enumeration from noisy multidimensional measurements even in scenarios where the number of components exceeds the size of the measurement tensor [7].

However, in CORCONDIA a proper user-defined threshold is required to ensure reliable performance. The optimal threshold increases with signal-to-noise ratio (SNR). In particular, for high SNRs, the CORCONDIA tends to over-enumeration by a couple of components under fixed threshold settings. To improve the detection accuracy, we propose an improved version of CORCONDIA by exploiting the reconstruction error that is readily available after PARAFAC decomposition. The motivation of using reconstruction error is that for high SNRs and at the neighborhood of the true number of components, the reconstruction error is more powerful at discriminating the correct number of components from incorrect ones than the core consistency.

Note that in the threshold-CORCONDIA [8], the difference of the core consistency for two adjacent candidate rank is used for enumeration instead. And the optimal threshold coefficients are searched in an R-D grid. The problem with this scheme is that a common threshold coefficient is used for a wide range of noise levels. Moreover, according to the definition the core consistency is unbounded in the negative direction, and hence using the difference seems not reasonable since a large gap between adjacent consistency values does not guarantee a high consistency value. As a result, the threshold-CORCONDIA [8] has poor performance with an empirical probability of correct detection (PoD) no more than 70% even at sufficiently high SNRs.

The notation used in this paper align with [9]. The r-mode vectors of a tensor $\boldsymbol{\mathcal{T}} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_R}$ are obtained by varying the r-th index within its range $(1, \ldots, I_r)$ and keeping all the other indices fixed. The r-mode unfolding of a tensor $\boldsymbol{\mathcal{T}}$, symbolized by $[\boldsymbol{\mathcal{T}}]_{(r)} \in \mathbb{C}^{I_r \times (I_1 \cdots I_{r-1} I_{r+1} \cdots I_R)}$, represents the matrix of r-mode vectors of $\boldsymbol{\mathcal{T}}$. The order of the columns is chosen in accordance with [9]. The r-mode product of $\boldsymbol{\mathcal{T}}$ and $\boldsymbol{U} \in \mathbb{C}^{J_r \times I_r}$ along the r-th mode is denoted as $\boldsymbol{\mathcal{T}} \times_r \boldsymbol{U} \in \mathbb{C}^{I_1 \times I_2 \cdots \times J_r \cdots \times I_R}$. It is obtained by multiplying the r-mode unfolding of $\boldsymbol{\mathcal{T}}$ from the left-hand side by $\boldsymbol{U}$. The operator $\mathrm{vec}(\cdot)$ converts a matrix or tensor into a vector by stacking its columns or 1-mode vectors on top of each other. The $\otimes$ and $\circ$ denote the Kronecker product and outer product, respectively. The superscripts $^{\mathrm{T}}$ and † represent matrix transpose and pseudo inverse, respectively.

The work described in this paper was supported by a grant from the NSFC/RGC Joint Research Scheme sponsored by the Research Grants Council of the Hong Kong and the National Natural Science Foundation of China (Project No.: N_CityU 104/11, 61110229)

2. DATA MODEL

The noise-free PARAFAC model is

$$\boldsymbol{\mathcal{X}} = \sum_{k=1}^{K} \boldsymbol{f}_k^{(1)} \circ \cdots \circ \boldsymbol{f}_k^{(R)}, \tag{1}$$

where $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{M_1 \times \cdots \times M_R}$ is the signal tensor, $\boldsymbol{f}_k^{(r)} = [f_k^{(r)}(1), \ldots, f_k^{(r)}(M_r)]^T$, $k = 1, \ldots, K$, $r = 1, \ldots, R$, is the k-th factor of the r-th mode.

By defining $\boldsymbol{F}^{(r)} = \left[\boldsymbol{f}_1^{(r)}, \ldots, \boldsymbol{f}_K^{(r)}\right] \in \mathbb{C}^{M_r \times K}$, $r = 1, \ldots, R$, (1) can be rewritten in terms of r-mode products as

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{I}}_{R,K} \times_1 \boldsymbol{F}^{(1)} \times_2 \boldsymbol{F}^{(2)} \cdots \times_R \boldsymbol{F}^{(R)}, \tag{2}$$

where $\boldsymbol{\mathcal{I}}_{R,K}$ represents the R-D identity tensor of size $K \times K \cdots \times K$, whose elements are equal to one when the indices $i_1 = i_2 = \cdots = i_R$ and zero otherwise. The $\boldsymbol{\mathcal{X}}$ is composed of the sum of K components each of which corresponds to a rank-1 tensor. The rank of a tensor is defined as the minimal number of rank-1 tensors that yields this tensor in a linear combination [9]. As in [3], we assume that the rank of $\boldsymbol{\mathcal{X}}$ is equal to the number of components, i.e., K.

In practice, the data are contaminated by noise and can be represented by

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{Z}}, \tag{3}$$

where $\boldsymbol{\mathcal{Z}}$ is the noise tensor collecting independent and identically distributed (i.i.d.) zero-mean circularly symmetric complex Gaussian (ZMCSCG) noise samples with variance of σ_z^2. The noise is assumed uncorrelated with the signals. Denoting $M = \prod_{r=1}^{R} M_r$, the SNR is defined as

$$\text{SNR} = \frac{\|\boldsymbol{\mathcal{X}}\|_F^2}{M\sigma_z^2}, \tag{4}$$

where $\|\cdot\|_F$ denotes the higher-order Frobenius norm of a tensor, which is defined as the square root of the sum of squared amplitude of its elements. Given the noisy measurement tensor $\boldsymbol{\mathcal{Y}}$, our goal is to estimate the number of components K.

3. REVIEW OF CORE CONSISTENCY DIAGNOSTIC

The principle behind the CORCONDIA is reviewed as follows. Given k as a candidate value for the number of components, the R factors are first estimated using the ALS [10]. Denote the resultant factor matrix estimates as $\left\{\hat{\boldsymbol{F}}^{(r)} | r = 1, \ldots, R\right\}$. We have

$$\boldsymbol{\mathcal{Y}} \simeq \hat{\boldsymbol{\mathcal{X}}} = \boldsymbol{\mathcal{I}}_{R,k} \times_1 \hat{\boldsymbol{F}}^{(1)} \times_2 \hat{\boldsymbol{F}}^{(2)} \cdots \times_R \hat{\boldsymbol{F}}^{(R)}. \tag{5}$$

Applying the vectorization on both sides of (5) yields

$$\text{vec}\,(\boldsymbol{\mathcal{Y}}) \simeq \left(\hat{\boldsymbol{F}}^{(R)} \otimes \cdots \otimes \hat{\boldsymbol{F}}^{(2)} \otimes \hat{\boldsymbol{F}}^{(1)}\right) \text{vec}(\boldsymbol{\mathcal{I}}_{R,k}). \tag{6}$$

Define a tensor $\boldsymbol{\mathcal{G}}_{R,k}$ of the same size as $\boldsymbol{\mathcal{I}}_{R,k}$ such that

$$\text{vec}(\boldsymbol{\mathcal{G}}_{R,k}) = \left(\hat{\boldsymbol{F}}^{(R)} \otimes \cdots \otimes \hat{\boldsymbol{F}}^{(2)} \otimes \hat{\boldsymbol{F}}^{(1)}\right)^{\dagger} \text{vec}\,(\boldsymbol{\mathcal{Y}})\,. \tag{7}$$

In the absence of noise and if the PARAFAC model is perfectly fulfilled with $k = K$, $\boldsymbol{\mathcal{G}}_{R,k}$ is equal to $\boldsymbol{\mathcal{I}}_{R,k}$. Otherwise, the closeness of $\boldsymbol{\mathcal{G}}_{R,k}$ to $\boldsymbol{\mathcal{I}}_{R,k}$, or more formally, the core consistency defined as [3]

$$\text{CC}(k) = 100\left(1 - \frac{\|\boldsymbol{\mathcal{G}}_{R,k} - \boldsymbol{\mathcal{I}}_{R,k}\|_F^2}{k}\right), \tag{8}$$

provides a measure of how well the k-component PARAFAC model fits the observations.

Formally, the estimate of CORCONDIA, denoted as $\hat{K}_{\text{CC}}$, is given by

$$\hat{K}_{\text{CC}} = \max\, k \,\text{subject to}\, \text{CC}(k) \geq \eta, \tag{9}$$

where $0 < \eta < 1.0$ is the threshold coefficient. An example of the core consistency profile is shown in Figure 1(a), where $M_1 = M_2 = M_3 = M_4 = 7$ and $K = 3$. We see that the CORCONDIA cannot discriminate between the underestimated and true numbers of components, since the former yields almost the same consistency as the true one. Therefore, it makes sense to take the signal number estimate as the highest valid number of components. Typically, $70\% \leq \eta \leq 90\%$ is used [3].

4. PROPOSED APPROACH: CORCONDIA AIDED BY RECONSTRUCTION ERROR

Other than underestimation, the CORCONDIA has low discrimination power in the neighborhood of the true number of components as well. In particular, for high SNRs, it occurs frequently that the overestimates by a couple of components yields a core consistency very close to 100% and clearly larger than 90%, as shown in Figure 1(a). Therefore, the CORCONDIA has a tendency to over-enumeration under fixed threshold settings. To handle this, we can set a higher threshold value. However, this will reduce the PoD for low-to-medium SNRs, where the true number of components often corresponds to a relatively small consistency value, as shown in Figure 1(b).

Therefore, the ideal threshold should increase with SNR. In order to achieve good detection performance at both low and high SNRs, we propose the following scheme. First, we adopt a threshold interval to obtain the lower and upper bounds of the estimates:

$$\hat{K}_{\text{lb}} = \max\, k \,\text{subject to}\, \text{CC}(k) \geq \eta_{\text{ub}}, \tag{10}$$

$$\hat{K}_{\text{ub}} = \max\, k \,\text{subject to}\, \text{CC}(k) \geq \eta_{\text{lb}}, \tag{11}$$

where η_{lb} and η_{ub} are the user-defined lower and upper bounds of the threshold coefficients. Empirically $5\% \leq \eta_{\text{lb}} \leq 25\%$ and $90\% \leq \eta_{\text{lb}} \leq 99\%$ work well.

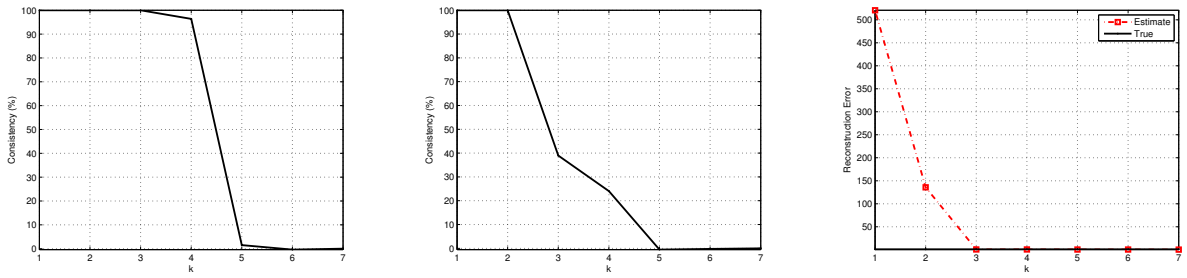
In general, the gap between $\hat{K}_{\text{lb}}$ and $\hat{K}_{\text{ub}}$ is small. To identify the final estimate, we exploit the reconstruction error which, for a candidate value $k > 0$, is defined as the sum of the squared differences between the measured and reconstructed k-component PARAFAC data:

$$r(k) = \left\|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{I}}_{R,k} \times_1 \hat{\boldsymbol{F}}^{(1)} \times_2 \hat{\boldsymbol{F}}^{(2)} \cdots \times_R \hat{\boldsymbol{F}}^{(R)}\right\|_F^2. \tag{12}$$

The motivation of using reconstruction error is that for high SNRs and at the neighborhood of the true number of components, the reconstruction error has more discriminative power than the core consistency. As shown in Figure 1(a), based on the core consistency we are apt to mistakenly choose the number of components as $k = 4$, while Figure 1(c) shows that using the reconstruction error, we can correctly identify the number of components as $k = 3$.

Denote the relative difference between the reconstruction error for $(k-1)$ and k as

$$D_r(k) = \frac{r(k-1) - r(k)}{r(k)}, \tag{13}$$



(a) CC: SNR = 30 dB (b) CC: SNR = 0 dB (c) RE: SNR = 30 dB

Fig. 1: Core consistency (CC) and reconstruction error (RE) versus k for high and low SNRs. $M_1 = M_2 = M_3 = M_4 = 7$. $K = 3$.

where $r(0) = \|\boldsymbol{\mathcal{Y}}\|_F^2$. We undertake a sequence of hypothesis tests starting with $\hat{K}_{\rm ub}$:

$$\begin{aligned} H_0 &: \ \hat{K}_{\rm ub} \text{ is NC }, D_r(\hat{K}_{\rm ub}) < \rho, \\ H_1 &: \ \hat{K}_{\rm ub} \text{ is PC }, D_r(\hat{K}_{\rm ub}) \geq \rho, \end{aligned} \tag{14}$$

where ρ is a pre-determined threshold, NC means noise component and PC means principal component. If $\hat{K}_{\rm ub}$ corresponds to NC, H_0 is accepted, and we proceed to the next hypothesis test with $\hat{K}_{\rm ub}-1$. We continue this way until the alternative hypothesis H_1 is accepted. If in all hypothesis tests from $\hat{K}_{\rm ub}$ to $\hat{K}_{\rm lb}$, H_0 is accepted, which is often the case for low SNRs less than 0 dB, then we choose $\hat{K}_{\rm ub}$ as the final estimate. The motivation of such a choice is that for low SNRs, the CORCONDIA has a tendency to underestimate the number of components. By choosing the estimation upper bound as the final estimate, this tendency can be alleviated and hence a higher PoD is expected.

Formally, defining

$$\mathfrak{K} = \left\{ \hat{K}_{\rm lb} \leq k \leq \hat{K}_{\rm ub} | D_r(k) \geq \rho \right\}, \tag{15}$$

the final estimate of the number of components is given by

$$\hat{K}_{\rm CC} = \begin{cases} \max\limits_{k \in \mathfrak{K}} k, & \mathfrak{K} \neq \emptyset; \\ \hat{K}_{\rm ub}, & \mathfrak{K} = \emptyset. \end{cases} \tag{16}$$

4.1. Threshold computation by Monte Carlo methods

The null hypothesis H_0 corresponds to over-enumeration by a certain number of signals not larger than $(K_{\rm ub} - K_{\rm lb})$. In the presence of strong signals (high SNRs), the probability of over-enumeration by δ components, $\delta = 1, \ldots, (K_{\rm ub} - K_{\rm lb})$, can be approximated by the probability of false alarm (Pfa) by the same number of components in noise-only measurements.

Denote the residual error of fitting the noisy measurement to a k-component PARAFAC model as σ_δ^2. Similarly with (13), we define the relative difference of adjacent fitting errors as

$$D_z(\delta) = \frac{\sigma_{\delta-1}^2 - \sigma_\delta^2}{\sigma_\delta^2}, \tag{17}$$

where $\sigma_0^2 = M\sigma_z^2$. It follows that the Pfa by δ components is

$$P_{\rm fa}(\delta) \simeq \Pr\left\{ D_z(\delta) \geq \rho(\delta) \right\} = E\left[1_{\mathfrak{D}} \left(\sigma_{\delta-1}^2, \sigma_\delta^2 \right) \right], \tag{18}$$

where

$$\mathfrak{D} = \left\{ \sigma_{\delta-1}^2 > \sigma_\delta^2 > 0 | D_r(\delta) \geq \rho(\delta) \right\}, \tag{19}$$

and $1_{\mathfrak{D}} \left(\sigma_{\delta-1}^2, \sigma_\delta^2 \right)$ is the indicator function equal to 1 for all elements in $\mathfrak{D}$ and 0 for all elements not in $\mathfrak{D}$.

The right-hand side of (18) can then be estimated by Monte Carlo simulations, which proceeds in the following steps:

1) Generate Q noise-only measurement tensor $\boldsymbol{\mathcal{Z}} \in \mathbb{C}^{M_1 \times \cdots \times M_R}$ of unit variance, where Q is the number of Monte Carlo runs. For the q-th, $q = 1, \ldots, Q$,
   a) Compute the reconstruction error $\sigma_{q,\delta}^2$ for $\delta = 1, 2, \ldots, \lceil \max(M_1, \ldots, M_R)/2 \rceil$ after PARAFAC decomposition, where $\lceil x \rceil$ denotes the smallest integer greater than or equal to x;
   b) Compute $D_z(\delta)$ and $1_{\mathfrak{D}} \left(\sigma_{q,\delta-1}^2, \sigma_{q,\delta}^2 \right)$ according to (17) and (18).
2) Estimate the Pfa by $\hat{P}_{\rm fa} = 1/Q \sum_{q=1}^{Q} 1_{\mathfrak{D}} \left(\sigma_{q,\delta-1}^2, \sigma_{q,\delta}^2 \right)$.

To obtain a relative estimation error of $\hat{P}_{\rm fa}$ that is smaller than α for a confidence level β, the number of Monte Carlo runs should satisfy $Q \geq c^2/(\alpha^2 P_{\rm fa})$ [11], where c is obtained from

$$\int_{-c}^{c} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{y^2}{2}\} dy = \beta.$$

In Figure 2 we have plotted the Pfa versus ρ. Here $\alpha = 0.1$, $\beta = 95\%$, and the minimum number of required Monte Carlo runs $Q = 38415$ is used.

From this curve, ρ is selected for each δ and for a given $P_{\rm fa}$. Considering that we do not know the extent of over-enumeration in H_0 of (14), we can take the maximum value of $\rho(\delta)$, $\delta = 1, \ldots, (K_{\rm ub} - K_{\rm lb})$ as the threshold ρ.

4.2. Note on complexity

The dominant computational load in the proposed scheme is the determination of the optimal threshold ρ via the Monte Carlo (MC) simulation in Section 4.1. Since the threshold is calculated through noise-only MC simulation and is a function of the measurement tensor size only, and in practical applications the measurement size is known and does not vary often, the MC simulation can be conducted offline. Besides, the reconstruction error in (12) is used in the stopping criterion for the iterative ALS PARAFAC decomposition algorithm, and therefore it is already known after the iteration stops and does not require extra computations as well.

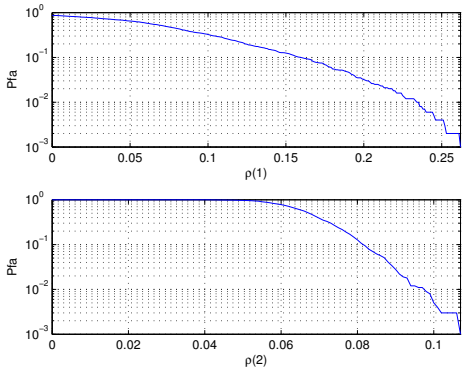


Fig. 2: Probability of false alarm versus $\rho(\delta)$ for threshold computation. $M_1 = M_2 = M_3 = M_4 = 5$.

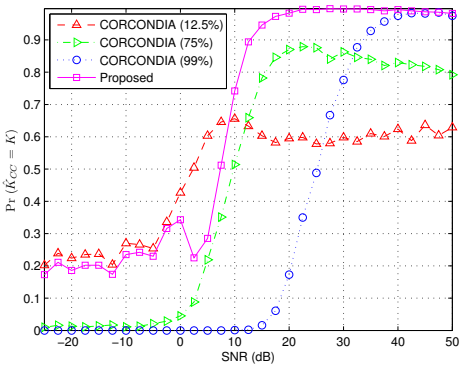


Fig. 3: PoD vs. SNR for a 3-D array of size $M_1 = 5, M_2 = 7, M_3 = 8$. $K = 4$.

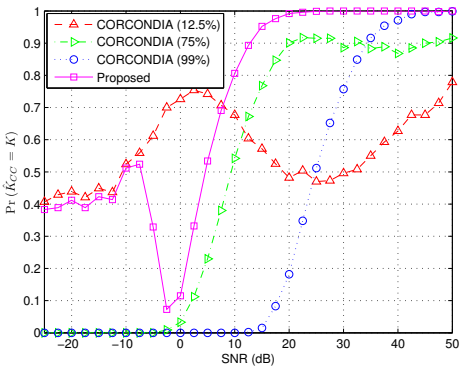


Fig. 4: PoD vs. SNR for a 4-D array of size $M_1 = M_2 = M_3 = M_4 = 5$. $K = 3$.

5. NUMERICAL EXAMPLES

In the PARAFAC model, the factor matrices contain i.i.d ZMCSCG entries with unit variance. The noise power σ_z^2 is scaled to obtain different SNRs. For each SNR, 1000 independent Monte Carlo runs have been conducted. The performance measure is the PoD, i.e., $\Pr(\hat{K}_{\rm CC} = K)$, averaged over noisy realizations of all Monte Carlo runs. In the proposed scheme, the lower and upper bounds of the threshold in the proposed scheme are chosen as $\eta_{\rm lb} = 12.5\%$ and $\eta_{\rm ub} = 99\%$. We compare our proposal with the following schemes: CORCONDIAs with a fixed threshold $\eta = 12.5\%, 99\%$ and the typical setting of $\eta = 75\%$.

First we consider a 3-D PARAFAC model of size $M_1 = 5, M_2 = 7, M_3 = 8$. In Figure 3, we plot the PoD versus SNR for $K = 4$ components. We see that the CORCONDIA with a fixed threshold $\eta = 12.5\%$ has a low PoD at high SNRs due to frequent over-enumeration, and the CORCONDIA with a fixed threshold $\eta = 99\%$ has a high PoD at sufficiently high SNRs but remarkably inferior performance at low-to-medium SNRs due to significant probability of under-enumeration. Our proposal combines the merits of both schemes, and significantly outperforms the scheme with the typical threshold of $\eta = 75\%$ at both low and high SNRs. Note that as the SNR decreases below 0 dB, the PoD rises to a level that is close to that of the CORCONDIA with a fixed threshold $\eta = 12.5\%$ instead of going down to zero. Similar observations are made in Figure 4. This is because for low SNRs, it frequently happens that H_1 is rejected in all hypothesis tests of candidate signal numbers within the interval $\left[\hat{K}_{\rm lb}, \hat{K}_{\rm ub}\right]$, in which case the candidate upper bound $\hat{K}_{\rm ub}$ is chosen as the final estimate according to (16). As mentioned in Section 4, such a strategy alleviates the tendency of the CORCONDIA to underestimation which results in a higher PoD.

In Figure 4, we use a 4-D PARAFAC model of size $M_1 = M_2 = M_3 = M_4 = 5$ for $K = 3$ components. Again, our proposal combines the advantages of the first scheme at low SNRs and the second scheme at high SNRs, and outperforms the last scheme with typical threshold of $\eta = 75\%$ for all SNRs.

6. CONCLUSION

The CORCONDIA is a conventional approach for determining the number of components in PARAFAC or Tucker 3 models. Its performance relies on a user-defined threshold whose optimal value varies with the SNR. In this work, we propose to first use a low threshold to obtain an upper bound of the estimate. The estimate takes the upper bound as its initial value, and is then gradually refined based on a sequence of hypothesis tests by exploiting the reconstruction error which is readily available after the PARAFAC decomposition. The proposed scheme results in accurate detection performance at both low and high SNRs, while almost no extra computation overhead is required.

REFERENCES

[1] R. Bro, "PARAFAC. Tutorial and applications," *Chemom. Intell. Lab. Syst.*, vol. 38, no. 2, pp. 149–171, Oct. 1997.

[2] N. M. Fabera, R. Bro, and P. K. Hopkec, "Recent developments in CANDECOMP/PARAFAC algorithms: A critical review," *Chemom. Intell. Lab. Syst.*, vol. 65, no. 1, pp. 119–137, Jan. 2003.

[3] R. Bro and H. A. L. Kiers, "A new efficient method for determining the number of components in PARAFAC models," *J. Chemom.*, vol. 17, pp. 274–286, 2003.

[4] D. Nion, K. N. Mokios, N. D. Sidiropoulos, and A. Potamianos, "Batch and adaptive PARAFAC-based blind separation of convolutive speech mixtures," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 6, pp. 1193 – 1207, 2010.

[5] C. A. R. Fernandes, G. Favier, and J. C. M. Mota, "PARAFAC-based channel estimation and data recovery in nonlinear MIMO spread spectrum communication systems," *Signal Processing*, vol. 91, no. 2, pp. 311–322, Feb. 2011.

[6] M. Kompany Zare, Y. Akhlaghi, and R. Bro, "Tucker core consistency for validation of restricted Tucker3 models," *Anal. Chim. Acta*, vol. 723, pp. 18–26, 2012.

[7] P. Comon, "Independent component analysis, a new concept?," *Signal Processing*, vol. 36, no. 3, pp. 287–314, April 1994.

[8] J. C. P. L. da Costa, M. Haardt, and F. Romer, "Robust methods based on the HOSVD for estimating the model order in PARAFAC models," in *Proc. 5th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM'08)*, Darmstadt, Germany, Jul. 2008, pp. 510–514.

[9] L. de Lathauwer, B. de Moor, and J. Vanderwalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253–1278, 2000.

[10] P. M. Kroonenberg and J. de Leeuw, "Principal component analysis of three-mode data by means of alternating least squares algorithms," *Psychometrika*, vol. 45, no. 1, pp. 69–97, Mar. 1980.

[11] A. Quinlan, J.-P. Barbot, P. Larzabal, and M. Haardt, "Model order selection for short data: An exponential fitting test (EFT)," *EURASIP Journal on Applied Signal Processing*, vol. Special Issue on Advances in Subspace-based Techniques for Signal Processing and Communications, 2007.