A MULTI-DIMENSIONAL MODEL ORDER SELECTION CRITERION WITH IMPROVED IDENTIFIABILITY

Kefei Liu†*, H.C. So*† *and Huang Lei*‡

† Department of Electronic Engineering, City University of Hong Kong, Hong Kong
‡ Department of Electronic and Information Engineering
Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China
E-mail: kefeilau@gmail.com, hcso@ee.cityu.edu.hk, lhuang@hitsz.edu.cn

ABSTRACT

A novel R-dimensional ($R \geq 3$) model order selection (MOS) criterion is proposed for estimating the number of sources embedded in noise. By extending the classical r-mode matrix unfolding of a Rth-order measurement tensor to multi-mode matrix unfolding, $(2^{R-1} - 1)$ unfolded matrices are obtained. To maximize the identifiability, the unfolded matrix whose number of rows is closest to that of the columns is chosen. Meanwhile, as the so-obtained unfolded matrix is of large size, a sequence of nested hypothesis tests on its associated eigenvalues is utilized for MOS in the framework of the random matrix theory. The maximum number of sources the proposed enumerator able to identify is on the order of the square root of the product of all dimension sizes, whereas the identifiability of existing criteria is limited to the maximum dimension size minus one. Numerical results are included to illustrate the performance of the proposed enumerator.

***Index Terms*—** model order selection, source enumeration, random matrix theory, tensor algebra

1. INTRODUCTION

R-dimensional (R-D) array signal processing where $R \geq 3$, have numerous applications such as wireless channel estimation, nuclear magnetic resonance (NMR) spectroscopy and multiple-input multiple-output (MIMO) radar imaging [1]. In these implementations, it is of considerable interest to accurately determine the source number by using the R-D structure of the array measurements. Usually, existing approaches require to stack the R-D measurements into a highly structured matrix, and then use one-dimensional (1-D) criteria for source enumeration. The 1-D source enumerators mainly consist of the minimum description length (MDL) [2], Akaike information criterion (AIC) [3], exponential fitting test (EFT) [4–6] and random matrix theory (RMT) based algorithm [7, 8]. In [5, 6, 9], a R-D extension of the 1-D MDL/AIC/EFT criteria is proposed, in which all r-mode ($r = 1, \cdots, R$) eigenvalues of the R-D measurement tensor $\boldsymbol{\mathcal{Y}} \in \mathbb{C}^{M_1 \times \cdots \times M_R}$ are calculated and combined to form global eigenvalues. As more sets of eigenvalues are employed for source enumeration, the R-D scheme is superior to the 1-D counterparts. However, their identifiability is only limited to $(\max(M_1, \cdots, M_R) - 1)$. To improve the identifiabiltiy, a novel RMT-based R-D source enumerator is proposed in this paper by exploiting the tensor structure of the observed data.

The work described in this paper was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project no. CityU 120911) and by the Natural Science Foundation of China (NSFC) under Grants 61171187 and 61161160564.

2. DATA MODEL

In general, the observations are modeled as

$$y_{(m_1,\cdots,m_R)} = \sum_{k=1}^{K} a_k^{(1)}(m_1) \cdots a_k^{(R)}(m_R) + z_{m_1,\cdots,m_R} \quad (1)$$

where $a_k^{(r)}(m_r)$ is the m_r-th element of the k-th factor of the r-th mode for $m_r = 1, \cdots, M_r$ ($r = 1, \cdots, R$), and $z_{m_1,\cdots,m_R}$ represents i.i.d. zero-mean circularly symmetric complex Gaussian (ZMCSCG) noise samples with variance of σ_z^2. The noise is uncorrelated both in all dimensions and with the signals. Usually, the R-th dimension corresponds to the temporal dimension, with $M_R = N$ denoting the number of snapshots and $a_k^{(R)}(m_R)$ denoting the complex amplitude of the k-th signal at the m_Rth time instant.

The tensor form of (1) is

$$\boldsymbol{\mathcal{Y}} = \sum_{k=1}^{K} \boldsymbol{a}_k^{(1)} \circ \cdots \circ \boldsymbol{a}_k^{(R)} + \boldsymbol{\mathcal{Z}} \quad (2)$$

where $\circ$ denotes the outer product, $\boldsymbol{\mathcal{Y}}$ is the noisy measurement tensor, $\boldsymbol{\mathcal{Z}} \in \mathbb{C}^{M_1 \times \cdots \times M_R}$ is the noise tensor collecting noise components, and $\boldsymbol{a}_k^{(r)} = [a_k^{(r)}(1), \cdots, a_k^{(r)}(M_r)]^T$. In the R-D harmonic retrieval model [1], $a_k^{(r)}$ ($r = 1, \cdots, R-1;\ k = 1, \cdots, K$) has a Vandermonde structure of the form of $a_k^{(r)} = [1, e^{j\mu_k^{(r)}}, \cdots, e^{j(M_r-1)\mu_k^{(r)}}]^T$, where $\{\mu_k^{(r)}\}$ represent the spatial frequencies.

The rank of a tensor is defined in [10]. In this paper, we assume that the rank of the noise-free measurement tensor $\boldsymbol{\mathcal{Y}}_0$ is equal to the number of signals, i.e., K. Given the noisy measurement tensor $\boldsymbol{\mathcal{Y}}$, our goal is to estimate the tensor rank, i.e., the number of signals K.

Let $M = \prod_{r=1}^{R} M_r$. The sample covariance matrix of the r-mode ($r = 1, \cdots, R$) matrix unfolding of $\boldsymbol{\mathcal{Y}}$ is defined as

$$\hat{\boldsymbol{R}}_{yy}^{(r)} = \frac{M_r}{M} \boldsymbol{\mathcal{Y}}_{(r)} \boldsymbol{\mathcal{Y}}_{(r)}^H \in \mathbb{C}^{M_r \times M_r} \quad (3)$$

where $\boldsymbol{\mathcal{Y}}_{(r)}$ is the r-mode matrix unfolding of $\boldsymbol{\mathcal{Y}}$ [10].

Classical eigenvalue-based detection criteria are 1-D based and cannot apply directly to the R-D measurement data. One solution is to convert the measurement tensor to matrix form by r-mode matrix unfolding, and then employ one or more sets of r-mode ($r =$

$1, \cdots, R$) eigenvalues for source enumeration [5, 6, 9]. This solution does not well exploit the inherent tensor structure of the measurement data, and as a result the identifiability is limited particularly when none of the dimension sizes is large enough.

3. ALGORITHM DEVELOPMENT

3.1. Random Matrix Theory [8]

For noise-only observations, the distribution of the largest sample eigenvalue under large-sample large-sensor asymptotic limit is characterized by Theorem 1 [11].

Theorem 1: Let $\boldsymbol{Y}$ be $p \times q \ (q \geq p)$ matrix with i.i.d. $\mathcal{CN}(0, \sigma^2 \boldsymbol{I}_p)$ entries. In the joint limit $p, q \rightarrow +\infty$, with $q/p \rightarrow \gamma \geq 1$, the distribution of the largest eigenvalue l_1 of the sample covariance matrix $\boldsymbol{R}_{yy} = \boldsymbol{Y}\boldsymbol{Y}^H/q$ converges to a Tracy-Widom distribution

$$\text{Pr}\left\{\frac{l_1/\sigma^2 - \tilde{\mu}_{q,p}}{\tilde{\sigma}_{q,p}}\right\} \rightarrow F_2(s) \tag{4}$$

where

$$\tilde{\mu}_{q,p} = \frac{1}{q}\left(\frac{1}{\sigma_{q-1,p}^{1/2}} + \frac{1}{\sigma_{q,p-1}^{1/2}}\right)\left(\frac{1}{\mu_{q-1,p}\,\sigma_{q-1,p}^{1/2}} + \frac{1}{\mu_{q,p-1}\,\sigma_{q,p-1}^{1/2}}\right)^{-1}$$

$$\tilde{\sigma}_{q,p} = \frac{1+\gamma_{q,p}}{q}\left(\frac{1}{\sigma_{q-1,p}} + \frac{\gamma_{q,p}}{\sigma_{q,p-1}}\right)^{-1}$$

with $\sigma_{q,p} = \left(\sqrt{q+1/2} + \sqrt{p+1/2}\right)\left(1/\sqrt{q+1/2} + 1/\sqrt{p+1/2}\right)^{1/3}$, $\mu_{q,p} = \left(\sqrt{q+1/2} + \sqrt{p+1/2}\right)^2$ and $\gamma_{q,p} = (\mu_{q-1,p}\,\sigma_{q,p-1}^{1/2})/(\mu_{q,p-1}\,\sigma_{q-1,p}^{1/2})$.

The above expressions provide $O(p^{-2/3})$ convergence rate in (4), and for finite q, p, they provide good approximations under the following two conditions

$$p \gg 1 \quad \text{and} \quad q/p \quad \text{is close to 1.} \tag{5}$$

For one or more signals embedded in noise, the distribution of the largest sample eigenvalue under large-sample large-sensor asymptotic limit is characterized by Theorem 2 [7].

Theorem 2: Consider a setting with K sufficiently strong signals. Then, in the asymptotic limit $p, q \rightarrow +\infty$, with $q/p \rightarrow \gamma > 0$, the $(K+1)$th largest sample eigenvalue has asymptotically the same Tracy-Widom distribution as the largest eigenvalue of a noise-only Wishart matrix, with parameters of $q, p-K$.

Based on Theorems 1 and 2, an accurate source enumerator has been developed in [8] where the signal number is estimated via a sequence of nested hypothesis tests. The algorithm works as follows. Denote by $\lambda_1 \geq \cdots \geq \lambda_p$ the eigenvalues of $\boldsymbol{R}_{yy}$. For $k = 1, \cdots, \min(p, q) - 1$, the hypothesis test is $\mathcal{H}_0$: only $k-1$ signals versus $\mathcal{H}_1$: at least k signals. $\mathcal{H}_0$ is rejected if $\lambda_k > \hat{\sigma}_z^2(k)\left(\tilde{\mu}_{q,p-k} + s(\alpha)\tilde{\sigma}_{q,p-k}\right)$ where $s(\alpha)$ is determined by $F_2(s(\alpha)) = 1 - \alpha$ with α being the confidence level. Here, $\hat{\sigma}_z^2(k)$ is an estimate for the unknown noise level σ_z^2 and is calculated based on the matrix perturbation theory and iteratively solving a non-linear system of equations, with remarkably improved accuracy over its maximum likelihood estimator [7, 8].

3.2. Proposed R-D Source Enumerator

First, the r-mode matrix unfolding of a tensor is extended to the multi-mode matrix unfolding.

Definition 1. Consider a J-D tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{I_1 \times \cdots \times I_J}$. For $1 \leq j_1 < \cdots < j_d \leq J \ (1 \leq d < J)$, the $(j_1, \cdots, j_d)$-mode matrix unfolding of $\boldsymbol{\mathcal{X}}$, denoted as $\boldsymbol{\mathcal{X}}_{(j_1,\cdots,j_d)} \in \mathbb{C}^{(I_{j_1}\cdots I_{j_d}) \times \frac{I_1 \cdots I_J}{I_{j_1}\cdots I_{j_d}}}$, is a matrix which contains the element $x_{(i_1,\cdots,i_J)}$ at the position with row number equals to

$$i_{j_1} + \cdots + (i_{j_d} - 1)I_{j_1} \cdots I_{j_{d-1}}$$

and column number equals to

$$\begin{aligned}
& i_1 + \cdots + (i_{j_1-1} - 1)I_1 \cdots I_{j_1-2} + (i_{j_1+1} - 1)I_1 \cdots I_{j_1-1} \\
& + \cdots + (i_{j_d-1} - 1)I_1 \cdots I_{j_1-1}I_{j_1+1} \cdots I_{j_d-2} + \\
& (i_{j_d+1} - 1)I_1 \cdots I_{j_1-1}I_{j_1+1} \cdots I_{j_d-1} + \cdots + \\
& (i_J - 1)I_1 \cdots I_{j_1-1}I_{j_1+1} \cdots I_{j_d-1}I_{j_d+1} \cdots I_{J-1}
\end{aligned}$$

Clearly, when $d = 1$, the multi-mode matrix unfolding reduces to the classical r-mode matrix unfolding. The sample covariance matrix associated with the $(r_1, \cdots, r_d)$-mode matrix unfolding $\boldsymbol{\mathcal{Y}}_{(r_1,\cdots,r_d)}$ is defined as

$$\hat{\boldsymbol{R}}_{yy}^{(r_1,\cdots,r_d)} = \frac{M_{r_1} \cdots M_{r_d}}{M} \boldsymbol{\mathcal{Y}}_{(r_1,\cdots,r_d)} \boldsymbol{\mathcal{Y}}_{(r_1,\cdots,r_d)}^H \tag{6}$$

where $\hat{\boldsymbol{R}}_{yy}^{(r_1,\cdots,r_d)} \in \mathbb{C}^{(M_{r_1}\cdots M_{r_d}) \times (M_{r_1}\cdots M_{r_d})}$. The set of eigenvalues of (6) is called the $(r_1, \cdots, r_d)$-mode eigenvalues of $\boldsymbol{\mathcal{Y}}$.

Note that each division of the index set $I = \{1, \cdots, R\}$ to two non-zero disjoint subsets $\{r_1, \cdots, r_d\}$ and $I/\{r_1, \cdots, r_d\}$ will result in a pair of mutually-transposed matrix unfolding: the $(r_1, \cdots, r_d)$-mode one with size of $(M_{r_1} \cdots M_{r_d}) \times \frac{M}{M_{r_1} \cdots M_{r_d}}$, and the $I/\{r_1, \cdots, r_d\}$-mode one with size of $\frac{M}{M_{r_1} \cdots M_{r_d}} \times (M_{r_1} \cdots M_{r_d})$. The total number of such division of I, the same as that of pairs of mutually-transposed matrix unfolding of $\boldsymbol{\mathcal{Y}}$, is

$$\left(C_R^1 + C_R^2 + \cdots + C_R^{R-1}\right)/2 = 2^{R-1} - 1 \tag{7}$$

As a pair of mutually-transposed unfolded matrices yield the same set of non-zero eigenvalues up to an irrelevant constant multiplication factor, we consider only the one whose number of rows is less than or equal to that of the columns. We choose from the resulting $2^{R-1} - 1$ matrices the one whose number of rows is closest to that of the columns for source enumeration, which is denoted as $\boldsymbol{Y}^{(1)}$.

Taking an example of $R = 4, M_1 = 5, M_2 = 3, M_3 = 6, M_4 = 4, \boldsymbol{\mathcal{Y}} \in \mathbb{C}^{5\times3\times6\times4}$, there are totally $2^{R-1} - 1 = 7$ unfolded matrices, namely, $\boldsymbol{\mathcal{Y}}_{(2,3)} = \boldsymbol{\mathcal{Y}}_{(1,4)}^T \in \mathbb{C}^{18\times20}$, $\boldsymbol{\mathcal{Y}}_{(1,2)} = \boldsymbol{\mathcal{Y}}_{(3,4)}^T \in \mathbb{C}^{15\times24}$, $\boldsymbol{\mathcal{Y}}_{(2,4)} = \boldsymbol{\mathcal{Y}}_{(1,3)}^T \in \mathbb{C}^{12\times30}$, $\boldsymbol{\mathcal{Y}}_{(3)} = \boldsymbol{\mathcal{Y}}_{(1,2,4)}^T \in \mathbb{C}^{6\times60}$, $\boldsymbol{\mathcal{Y}}_{(1)} = \boldsymbol{\mathcal{Y}}_{(2,3,4)}^T \in \mathbb{C}^{5\times72}$, $\boldsymbol{\mathcal{Y}}_{(4)} = \boldsymbol{\mathcal{Y}}_{(1,2,3)}^T \in \mathbb{C}^{4\times90}$, and $\boldsymbol{\mathcal{Y}}_{(2)} = \boldsymbol{\mathcal{Y}}_{(1,3,4)}^T \in \mathbb{C}^{3\times120}$. Hence $\boldsymbol{Y}^{(1)} = \boldsymbol{\mathcal{Y}}_{(2,3)} \in \mathbb{C}^{18\times20}$.

The RMT-based source enumerator is originally proposed to estimate the number of 1-D signals. For estimating the number of R-D signals at hand, we propose to apply the algorithm to the set of eigenvalues associated with $\boldsymbol{Y}^{(1)}$, and call the resultant criterion R-D RMT. This R-D RMT criterion has the following advantages:

I) Higher detection probability. From the choice of $\boldsymbol{Y}^{(1)}$, we can see that among all unfolded matrices, the difference between the number of rows P_1 and that of columns Q_1 of $\boldsymbol{Y}^{(1)}$ is the smallest. Therefore, conditions (5) are better satisfied and better approximations can be achieved by using the Tracy-Widom distribution.

II) Improved identifiability. Using $\boldsymbol{Y}^{(1)}$, the R-D RMT criterion can detect up to $(P_1 - 1)$ signals, while the R-D criterion in [9] can detect at most $(\max(M_1, \cdots, M_R) - 1)$ signals. For $R > 3$, $P_1 - 1 \simeq \sqrt{M_1 \times \cdots \times M_R} \gg \max(M_1, \cdots, M_R) - 1$.

In the absence of noise, it can be proved that the rank of $\boldsymbol{Y}^{(1)}$ is less than or equal to the rank of $\boldsymbol{\mathcal{Y}}$, i.e., the number of signals. Furthermore, for the R-D harmonic retrieval model, when $K \leq P_1$, the rank of $\boldsymbol{Y}_0^{(1)}$ has been shown by extensive numerical simulations to be equal to the rank of $\boldsymbol{\mathcal{Y}}_0$ almost surely, if $\{\mu_k^{(r)} | r = 1, \cdots, R-1;\ k = 1, \cdots, K\}$ are drawn from a continuous distribution, and the source amplitudes are i.i.d Gaussian distributed.

The performance of the RMT-based source enumerator with constant confidence level is not robust to the number of signals. For a relatively low confidence level, it works well when the number of signals is small, while tends to underestimate the number of signals when the source number is large. For a relatively high confidence level, on the contrary, it works well for large source numbers, while tends to overestimation for small source numbers. To overcome this drawback, we propose to use an adaptive confidence level instead of constant one in the R-D RMT criterion. Mathematically, in the kth $(k = 1, \cdots, P_1 - 1)$ test,

$$\alpha(k) = \begin{cases} \exp\left[\log \alpha_1 + (k-1)/(c-1) \cdot \log(\alpha_{\text{mid}}/\alpha_1)\right], & k \leq c \\ (\alpha_1 + \alpha_2) - \alpha(P_1 - k), & \text{otherwise} \end{cases} \tag{8}$$

where $c = \lfloor P_1/2 \rfloor$ denotes the largest integer less than or equal to $P_1/2$, α_1 and α_2 are the user-defined lower and upper bounds of α, respectively, and $\alpha_{\text{mid}} = (\alpha_1 + \alpha_2)/2$. The resultant source enumerator is called adaptive R-D RMT.

4. NUMERICAL EXAMPLES

We take the R-D harmonic retrieval model for illustration. The spatial frequencies $\mu_k^{(r)} (r = 1, \cdots, R-1;\ k = 1, \cdots, K)$ are i.i.d uniformly distributed within $[-\pi, \pi]$, and the sources are i.i.d ZMCSCG distributed with equal power σ_s^2. The SNR is defined as SNR $= \sigma_s^2/\sigma_z^2$. 2000 independent Monte Carlo runs have been conducted. The performance measure is the probability of detection (PoD), i.e., Pr$(\hat{K} = K)$, averaged over spatial frequencies, sources and noise realizations of all Monte Carlo runs.

The PoD's of the following schemes are compared: MDL/AIC using eigenvalues associated with $\boldsymbol{Y}^{(1)}$, two R-D RMT versions with constant confidence levels $\alpha_1 = 10^{-4}$ and $\alpha_2 = 0.25$, and adaptive RMT with $\alpha(k)$ defined as in (8). Figure 1 shows the curves of $\alpha(k)$ used in the following simulation. Note that all investigated schemes are new because they employ $\boldsymbol{Y}^{(1)}$.

First we consider a system where $K = 5$ and $K = 54$ sources impinge on a 3-D array of size $M_1 = 9, M_2 = 8, M_3 = 7$, with $M_4 = N = 10$ snapshots. In Figure 2, the PoD of the RMT scheme using $\boldsymbol{Y}^{(1)}$ and other unfolded matrices for $\alpha_1 = 10^{-4}$ and $K = 5$ are compared. It can be seen that the PoD of the RMT scheme using $\boldsymbol{Y}^{(1)}$ is always better than or comparable to that using other unfolded matrices. Similar observations are obtained for other parameter settings.

Figure 3 compares the PoD of different schemes. We see that the R-D RMT scheme can detect the true number of signals in both source number scenarios. Compared with the R-D MDL/AIC/EFT schemes [9] which can detect at most 9 sources, the R-D RMT identifiability is significantly enhanced. However, the performance of the R-D RMT with constant confidence level is not robust against the case of wide-range source number variation. For small number of signals, R-D RMT with $\alpha_1 = 10^{-4}$ works well and is consistent as the SNR goes to infinity, while for large number of signals, it tends to underestimate the true number of signals and when $K = 54$, the PoD is only 0.6 even at high SNR. On the contrary, R-D RMT with $\alpha_2 = 0.25$ tends to overestimation for small source numbers and the PoD is less than 0.8 even at high SNR when $K = 5$, while for large source numbers, it works well and gains large-SNR consistency. The adaptive R-D RMT is a good compromise of both versions and is robust against source number variation because it works well for both small and large source numbers and achieves near consistency in the larger-SNR limit. For relatively large source numbers, AIC/MDL is not consistent when the SNR goes to infinity, having a tendency to overestimate the true number of signals. This is consistent with [12].

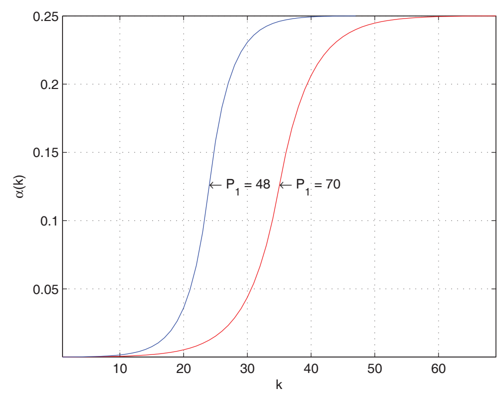

Fig. 1: Confidence level $\alpha(k)$ as a function of hypothesized number of signals k when $[\alpha_1, \alpha_2] = [10^{-4}, 0.25]$.

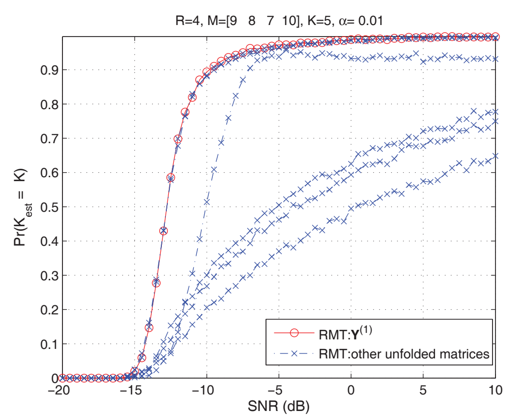

Fig. 2: Comparison of PoD of RMT by using $\boldsymbol{Y}^{(1)}$ and other unfolded matrices ($\alpha = 10^{-4}, K = 5$).

Figure 4 considers a system with a 4-D array of size $M_1 = 3, M_2 = 4, M_3 = 5, M_4 = 6$, with $M_5 = N = 8$ snapshots. In this case, the R-D EFT/MDL/AIC schemes proposed in [9] can detect at most 7 sources, while the proposed R-D RMT scheme can detect up to 40. And again, the adaptive RMT scheme shows good robustness in a wide range of number of signals and its performance outperforms the AIC/MDL.

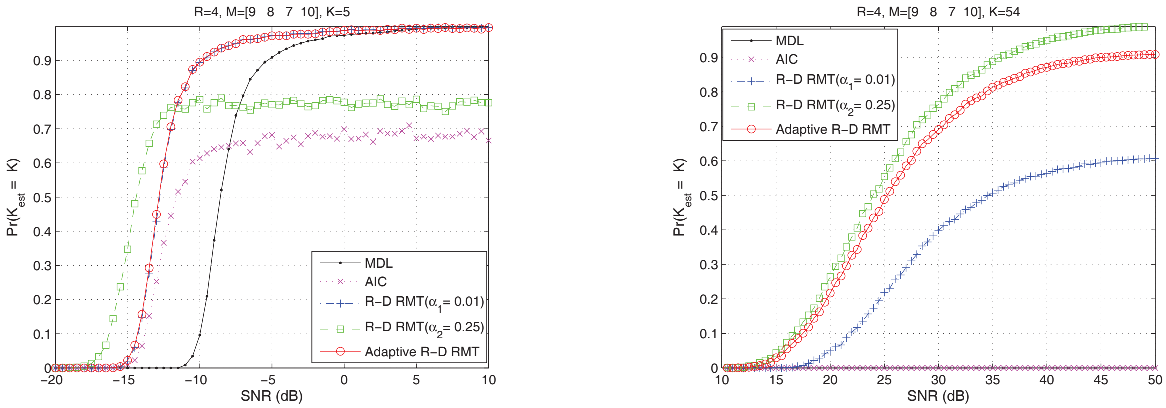

Fig. 3: PoD vs. SNR for a 3-D array of size $M_1 = 9, M_2 = 8, M_3 = 7$, with $M_4 = 10$ (left: $K = 5$; right: $K = 54$).

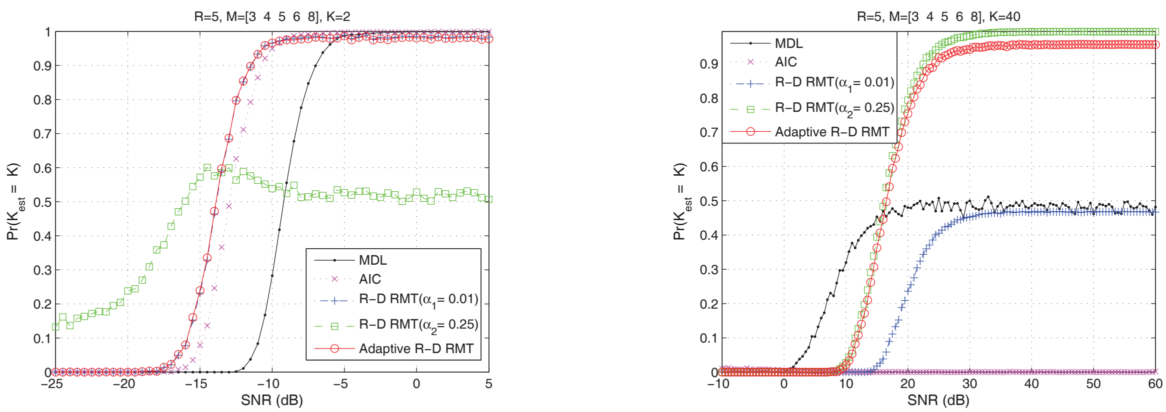

Fig. 4: PoD vs. SNR for a 4-D array of size $M_1 = 3, M_2 = 4, M_3 = 5, M_4 = 6$, with $M_5 = 8$ snapshots (left: $K = 2$; right: $K = 40$).

5. CONCLUSION

We have proposed a novel R-D ($R \geq 3$) detection criterion for accurate source enumeration. With the generalized multi-mode matrix unfolding of the measurement tensor and the random matrix theory (RMT), the eigenvalues associated with the unfolded matrix whose number of rows is closest to that of the columns are used for model order selection via a sequence of nested hypothesis tests. Compared with existing R-D detection schemes, the proposed R-D RMT criterion is able to significantly improve the identifiability. The adaptive version of the R-D RMT has robust performance and works well in a wide range of number of signals.

REFERENCES

[1] D. Nion and N. D. Sidiropoulos, "Tensor algebra and multidimensional harmonic retrieval in signal processing for MIMO radar," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5693–5705, 2010.

[2] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, pp. 465–471, 1979.

[3] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. 19, no. 6, pp. 716–723, 1974.

[4] 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*, pp. 54–64, 2007.

[5] J. P. C. L. da Costa, M. Haardt, F. Romer, and G. Del Galdo, "Enhanced model order estimation using higher-order arrays," in *Proc. Conf. Record of the Forty-First Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, USA, Nov. 2007, pp. 412–416.

[6] J. P. C. L. da Costa, A. Thakre, F. Roemer, and M. Haardt, "Comparison of model order selection techniques for high-resolution parameter estimation algorithms," in *Proc. 54th International Scientific Colloquium (IWK)*, Ilmenau, Germany, Sept. 2009.

[7] S. Kritchman and B. Nadler, "Determining the number of components in a factor model from limited noisy data," *Chem. Int. Lab. Syst.*, vol. 94, pp. 19–32, 2008.

[8] S. Kritchman and B. Nadler, "Non-parametric detection of the number of signals: Hypothesis testing and random matrix theory," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3930–3941, 2009.

[9] J. P. C. L. da Costa, F. Roemer, M. Haardt, and R. T. de Sousa Jr., "Multi-dimensional model order selection," *Journal on Advances in Signal Processing*, vol. 26, 2011.

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

[11] N. El Karoui, "A rate of convergence result for the largest eigenvalue of complex white Wishart matrices," *Ann. Prob.*, vol. vol. 36, no. 6, pp. 2077–2117, 2006.

[12] Q. Ding and S. Kay, "Inconsistency of the MDL: On the performance of model order selection criteria with increasing signal-to-noise ratio," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 1959–1969, 2011.