Alignment and matching tests for high-dimensional tensor signals via tensor contraction

Jianfeng Yao

School of Data Science
The Chinese University of Hong Kong (Shenzhen)
e-mail: jeffyao@cuhk.edu.cn

This is a joint work with Ruihan Liu and Zhenggang Wang,
Department of Statistics and Actuarial Science,
The University of Hong Kong

MSC2020 subject classifications: Primary 62H15; secondary 60B20,62H10.
Keywords and phrases: High-dimensional tensors, Low-rank tensors, Tensor signal alignment, Tensor signal matching, Tensor contraction, Linear spectral statistics, Random matrix theory.

In the era of "big data", the analysis of high-dimensional tensor data has become increasingly important in various fields, including genomics, economics, image analysis, and machine learning. High-order tensor data often exhibit intrinsic low-rank structures [14, 25]. To capture these low-rank structures, the "signal plus noise" tensor model has been widely adopted [9, 11, 15]. Let $n_1, \dots, n_d \in \mathbb{N}^+$ denote d dimension numbers, where $d \geq 3$, and let $N = n_1 + \cdots + n_d$. The d-fold rank-R spiked tensor model is defined as:

$$\boldsymbol{T} = \sum_{r=1}^{R} \beta_r \boldsymbol{x}^{(r,1)} \otimes \cdots \otimes \boldsymbol{x}^{(r,d)} + \frac{1}{\sqrt{N}} \boldsymbol{X}, \tag{1}$$

where $\beta_1 \geq \cdots \geq \beta_R > 0$ are the signal-to-noise ratios (SNRs), $\{\boldsymbol{x}^{(1,l)}, \cdots, \boldsymbol{x}^{(R,l)}\}$ are mutually orthogonal unit vectors $\mathbb{R}^{n_l}$ for each $1 \leq l \leq d$ [13], and $\boldsymbol{X} = (X_{i_1 \cdots i_d})_{n_1 \times \cdots \times n_d} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ is a noise tensor with independent and identically distributed (i.i.d.) entries, each having zero mean and unit variance. Specifically, the rank-1 spiked tensor model [21] is given by:

$$\boldsymbol{T} = \beta \boldsymbol{x}^{(1)} \otimes \cdots \otimes \boldsymbol{x}^{(d)} + \frac{1}{\sqrt{N}} \boldsymbol{X}, \tag{2}$$

where $\beta > 0$ is the single SNR of the model.

The primary focus of most existing literature is on recovering the signal vectors $\{\boldsymbol{x}^{(1,l)}, \dots, \boldsymbol{x}^{(R,l)}\}$, $1 \leq l \leq d$ from the observed tensor $\boldsymbol{T}$, with a particular emphasis on the computational efficiency of recovery algorithms. In the case of the rank-one model (2) with symmetric and i.i.d. Gaussian noise $\boldsymbol{X}$, [10] showed that computing the maximum likelihood (ML) estimator of $\beta \boldsymbol{x}^{(1)} \otimes \cdots \otimes \boldsymbol{x}^{(d)}$ is in general NP-hard, and [1] provided a comprehensive discussion on the relationship between the computational complexity of the ML estimator and the value of the SNR β. To reduce the computational complexity, [21] proposed the use of the power iteration method and approximate message passing (AMP) algorithms. These two methods have been extensively investigated by [5, 7, 12, 15, 20] for AMP and by [11] for power iteration. Moreover, [21] introduced the tensor unfolding method, which involves unfolding the tensor data $\boldsymbol{T}$ into matrices, enabling the recovery of signals through Principal Component Analysis (PCA). [6] conducted a comprehensive study of the tensor unfolding method for the general asymmetric model (2) under fairly general noise distribution assumptions.

However, when the SNRs fall below the phase transition threshold, these recovery methods often fail. In such cases, a less ambitious but potentially more achievable goal is to test the alignment of a signal in $\boldsymbol{T}$ with a given directional tensor $\boldsymbol{a}^{(1)} \otimes \cdots \otimes \boldsymbol{a}^{(d)}$, where $\boldsymbol{a}^{(j)}, 1 \leq j \leq d$ are d given directional unit vectors in $\mathbb{R}^{n_j}$, respectively. This leads to the following *tensor signal alignment test* between two hypotheses:

$$\begin{aligned} &H_0 : \boldsymbol{a}^{(l)} \perp \boldsymbol{x}^{(r,l)} \quad \text{for } 1 \leq l \leq d,\ 1 \leq r \leq R. \\ &H_1 : \text{there exists at least one } 1 \leq l \leq d,\ 1 \leq r \leq R \text{ such that } \boldsymbol{a}^{(\mathrm{l})} \not\perp \boldsymbol{x}^{(\mathrm{r,l})}. \end{aligned} \tag{3}$$

Despite the tensor signal alignment test appearing more tractable than signal recovery, to the best of our knowledge, there is no established and rigorously justified procedure for addressing this problem. The difficulty stems from the high dimensionality of the tensors and the lack of a meaningful test statistic.

We leverage the tensor contraction operator $\boldsymbol{\Phi}_d$, originally proposed in [22], which maps an arbitrary tensor $\boldsymbol{T}$ and unit vectors $\{\boldsymbol{a}^{(j)}\}$ to a matrix $\boldsymbol{R}$:

$$\begin{aligned} &\boldsymbol{\Phi}_d : \mathbb{R}^{n_1 \times \cdots \times n_d} \times \mathbb{S}^{n_1 - 1} \times \cdots \times \mathbb{S}^{n_d - 1} \longrightarrow \mathbb{R}^{N \times N}, \\ &(\boldsymbol{T}, \boldsymbol{a}^{(1)}, \cdots, \boldsymbol{a}^{(d)}) \longmapsto \boldsymbol{R} = \begin{pmatrix} \boldsymbol{0}_{n_1 \times n_1} & \boldsymbol{T}^{12} & \cdots & \boldsymbol{T}^{1d} \\ (\boldsymbol{T}^{12})' & \boldsymbol{0}_{n_2 \times n_2} & \cdots & \boldsymbol{T}^{2d} \\ \vdots & \vdots & \ddots & \vdots \\ (\boldsymbol{T}^{1d})' & (\boldsymbol{T}^{2d})' & \cdots & \boldsymbol{0}_{n_d \times n_d} \end{pmatrix}. \end{aligned} \tag{4}$$

Here, for a pair of indices $1 \leq j_1 < j_2 \leq d$, $\boldsymbol{T}^{j_1 j_2}$ is an $n_{j_1} \times n_{j_2}$ matrix, called *second order contraction matrix of* $\boldsymbol{T}$ *along the directions* $\{\boldsymbol{a}^{(j_1)}, \boldsymbol{a}^{(j_2)}\}$, as introduced in [16]. It is defined by:

$$\boldsymbol{T}^{j_1 j_2} = \left[\sum_{i_j = 1, j \neq j_1, j_2}^{n_j} T_{i_1 \cdots i_d} \prod_{l=1, l \neq j_1, j_2}^{d} a_{i_l}^{(l)} \right]_{n_{j_1} \times n_{j_2}}. \tag{5}$$

From a mathematical perspective, the contraction operator $\mathbf{\Phi}_d$ has several advantages. Firstly, $\mathbf{\Phi}_d$ is linear in $\boldsymbol{T}$. When applied to the R-rank tensor in (1), we have

$$\begin{aligned} \boldsymbol{R} &= \mathbf{\Phi}_d(\boldsymbol{T}, \boldsymbol{a}^{(1)}, \cdots, \boldsymbol{a}^{(d)}) \\ &= \sum_{r=1}^{d} \beta_r \mathbf{\Phi}_d(\boldsymbol{x}^{(r,1)} \otimes \cdots \otimes \boldsymbol{x}^{(r,d)}, \boldsymbol{a}^{(1)}, \cdots, \boldsymbol{a}^{(d)}) + \frac{1}{\sqrt{N}} \mathbf{\Phi}_d(\boldsymbol{X}, \boldsymbol{a}^{(1)}, \cdots, \boldsymbol{a}^{(d)}), \\ &= \boldsymbol{S} + \boldsymbol{M}. \end{aligned} \tag{6}$$

where $\boldsymbol{S}$ is the contracted signal matrix containing the $\boldsymbol{R}$ tensor signals, and $\boldsymbol{M}$ is the residual matrix representing pure noise. Under the null hypothesis H_0, $\boldsymbol{S} = \mathbf{0}$ implying $\boldsymbol{R} = \boldsymbol{M}$. In contrast, under the alternative H_1, $\boldsymbol{S} \neq \mathbf{0}$, result in $\boldsymbol{R} \neq \boldsymbol{M}$.

Furthermore, both the contracted signal matrix $\boldsymbol{S}$ and noise matrix $\boldsymbol{M}$ are symmetric, with $\boldsymbol{S}$ having a finite rank. This allows us to analyze the contracted data matrix $\boldsymbol{R}$ using linear spectral statistics (LSS), a powerful tool from random matrix theory. Central limit theorems for LSS of random matrices have received much attention in high-dimensional statistics, see [2, 3, 17, 19, 26] for a few classical references. In our case, by employing an appropriate LSS of $\boldsymbol{R}$ with an established asymptotic distribution, we can effectively distinguish between the two hypotheses.

We first establish that the eigenvalue distribution of $\boldsymbol{R}$ has a limit ν when the d dimensions $\{n_j\}$ grow to infinity in comparable rates. Next, we introduce the following test statistic:

$$\widehat{T}_N^{(d)} = \|\boldsymbol{R}\|_2^2 - N \int_{-\infty}^{\infty} x^2 \nu(dx). \tag{7}$$

Here, $\|\boldsymbol{R}\|_2^2 = \sum_{i,j=1}^{N} R_{i,j}^2$ is a linear spectral statistic of $\boldsymbol{R}$. As one of the main results of this paper, we establish that under the null hypothesis H_0,

$$\frac{\widehat{T}_N^{(d)} - \xi_N^{(d)}}{\sigma_N^{(d)}} \xrightarrow{d} \mathcal{N}(0,1), \tag{8}$$

where $\xi_N^{(d)}$ and $\sigma_N^{(d)}$ are known parameters that can be calculated numerically. Under the alternative hypothesis H_1,

$$\frac{\widehat{T}_N^{(d)} - \xi_N^{(d)}}{\sigma_N^{(d)}} - \mathcal{D}^{(d)}/\sigma_N^{(d)} \xrightarrow{d} \mathcal{N}(0,1), \tag{9}$$

where $\mathcal{D}^{(d)}/\sigma_N^{(d)}$ is a positive mean drift. Consequently, the asymptotic normal distribution in (8) enables us to construct a test for a given significance level α, while the distribution in (9) guarantees a positive power for the test, which depends on the magnitude of $\mathcal{D}^{(d)}/\sigma_N^{(d)}$.

When $d = 2$, the tensor model (1) reduces to a finite-rank perturbed or spiked random matrix. In this context, the signal alignment test in (3) can be seen as a tensor extension of existing tests for the presence of spikes along given directions, as studied by [4, 8, 18, 23, 24].

However, when $d \geq 3$, a fundamental difference emerges: the elements $\boldsymbol{T}^{j_1 j_2}$ in the contracted data matrix $\boldsymbol{R}$ become correlated. This correlation significantly increases the complexity of studying the matrix, making the analysis more challenging compared to the $d = 2$ case. The presence of these correlations necessitates the development of novel techniques to effectively analyze the eigenvalue distribution and establish the asymptotic properties of the test statistic $\widehat{T}_N^{(d)}$ in high dimensions.

The main contributions of this article are as follows.

(i) We conduct an in-depth analysis of the contracted data matrix $\boldsymbol{R}$, whose entries display significant correlations and deviate from traditional random matrix models in which the elements of the noise matrix are typically assumed to be independent of one another, including

(a) The characterization of its limiting spectral distribution (LSD) through a vector Dyson equation, along with entrywise behaviors of the resolvent.

(b) The establishment of CLT for a broad class of its LSS.

(ii) We establish a rigorous procedure for the tensor signal alignment test (3) by establishing the normality asymptotic of the test statistic and deriving its power function under a general alternative hypothesis.

(iii) We also address the problem of testing for the matching of two high-dimensional low-rank tensor signals. To tackle this problem, we employ an approach similar to the one established for the tensor signal alignment test.

The contributions presented in this article are novel. One notable innovation is that our tensor signal model in (1) allows for non-Gaussian and non-symmetric signals. This sets our work apart from most existing literature on high-dimensional tensor data models, which typically assumes symmetry or Gaussianity for either the tensor signal, the tensor noise, or both.

References

[1] Arous, G. B., Mei, S., Montanari, A. and Nica, M. (2019). The landscape of the spiked tensor model. *Communications on Pure and Applied Mathematics* **72** 2282–2330.

[2] Bai, Z., Jiang, D., Yao, J. F. and Zheng, S. (2009). Corrections to LRT on large-dimensional covariance matrix by RMT. *The Annals of Statistics* **37** 3822-3840.

[3] Bai, Z. and Silverstein, J. W. (2004). CLT for Linear Spectral Statistics of Large-Dimensional Sample Covariance Matrices. *Annals of Probability* 553–605.

[4] Bao, Z., Ding, X., Wang, J. and Wang, K. (2022). Statistical inference for principal components of spiked covariance matrices. *The Annals of Statistics* **50** 1144–1169.
[5] Ben Arous, G., Gheissari, R. and Jagannath, A. (2020). Algorithmic thresholds for tensor PCA. *Annals of Probability* **48** 2052–2087.
[6] Ben Arous, G., Huang, D. Z. and Huang, J. (2023). Long Random Matrices and Tensor Unfolding. *The Annals of Applied Probability* **33** 5753–5780.
[7] Chen, W.-K. (2019). Phase transition in the spiked random tensor with Rademacher prior. *The Annals of Statistics* **47** 2734–2756.
[8] Hallin, M., Paindaveine, D. and Verdebout, T. (2010). Optimal rank-based testing for principal components. *Annals of statistics* **38** 3245–3299.
[9] Han, Y. and Zhang, C.-H. (2022). Tensor principal component analysis in high dimensional CP models. *IEEE Transactions on Information Theory* **69** 1147–1167.
[10] Hillar, C. J. and Lim, L.-H. (2013). Most tensor problems are NP-hard. *Journal of the ACM (JACM)* **60** 1–39.
[11] Huang, J., Huang, D. Z., Yang, Q. and Cheng, G. (2022). Power iteration for tensor PCA. *Journal of Machine Learning Research* **23** 1–47.
[12] Jagannath, A., Lopatto, P. and Miolane, L. (2020). Statistical thresholds for tensor PCA. *The Annals of Applied Probability* **30** 1910–1933.
[13] Kolda, T. G. (2001). Orthogonal tensor decompositions. *SIAM Journal on Matrix Analysis and Applications* **23** 243–255.
[14] Kolda, T. G. and Bader, B. W. (2009). Tensor decompositions and applications. *SIAM review* **51** 455–500.
[15] Lesieur, T., Miolane, L., Lelarge, M., Krzakala, F. and Zdeborová, L. (2017). Statistical and computational phase transitions in spiked tensor estimation. *2017 IEEE International Symposium on Information Theory (ISIT)* 511–515.
[16] Lim, L.-H. (2005). Singular values and eigenvalues of tensors: a variational approach. In *1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, 2005.* 129–132. IEEE.
[17] Lytova, A. and Pastur, L. (2009). Central limit theorem for linear eigenvalue statistics of random matrices with independent entries. *The Annals of Probability* **37** 1778 – 1840.
[18] Naumov, A., Spokoiny, V. and Ulyanov, V. (2019). Bootstrap confidence sets for spectral projectors of sample covariance. *Probability Theory and Related Fields* **174** 1091–1132.
[19] Pan, G. M. and Zhou, W. (2008). Central limit theorem for signal-to-interference ratio of reduced rank linear receiver. *Ann. Appl. Probab.* 1232-1270.
[20] Perry, A., Wein, A. S. and Bandeira, A. S. (2020). Statistical limits of spiked tensor models. *Annales de l'Institut Henri Poincaré. Probabilités et Statistiques* **56** 230–264.
[21] Richard, E. and Montanari, A. (2014). A statistical model for tensor

PCA. *Advances in Neural Information Processing Systems* **27**.

[22] Seddik, M. E. A., Guillaud, M. and Couillet, R. (2024). When random tensors meet random matrices. *Annals of Applied Probability* **34** 203 – 248.

[23] Silin, I. and Fan, J. (2020). Hypothesis testing for eigenspaces of covariance matrix. *arXiv preprint arXiv:2002.09810*.

[24] Silin, I. and Spokoiny, V. (2018). Bayesian inference for spectral projectors of the covariance matrix. *Electronic Journal of Statistics* **12** 1948–1987.

[25] Udell, M. and Townsend, A. (2019). Why are big data matrices approximately low rank? *SIAM Journal on Mathematics of Data Science* **1** 144–160.

[26] Zheng, S., Bai, Z. and Yao, J. (2017). CLT for eigenvalue statistics of large-dimensional general Fisher matrices with applications. *Bernoulli* **23** 1130-1178.