Truncated HOSVD: A Random Matrix Analysis

Hugo Lebeau
hugolebeau.github.io

October 2, 2024
Random Tensors and Related Topics

Spiked Random Tensor Model

$$\mathfrak{T} = \mathfrak{P} + \frac{1}{\sqrt{N}}\mathfrak{N}$$

Spiked Random Tensor Model

$$\mathfrak{T} = \qquad \mathcal{P} \qquad + \quad \frac{1}{\sqrt{N}}\mathfrak{N} \quad \in \mathbb{R}^{n_1 \times n_2 \times n_3}, \qquad \begin{array}{l} N = n_1 + n_2 + n_3 \\ \mathfrak{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1) \end{array}$$

Spiked Random Tensor Model

$$\mathcal{T} = \underbrace{\mathcal{P}}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$$N = n_1 + n_2 + n_3$$

$$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

Spiked Random Tensor Model

$$\mathcal{T} = \underbrace{\mathcal{P}}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$N = n_1 + n_2 + n_3$

$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$

Goal

Estimate $\mathcal{P}$ with a low-rank approximation of $\mathcal{T}$.

Spiked Random Tensor Model

$$\mathcal{T} = \underbrace{\mathcal{P}}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$$N = n_1 + n_2 + n_3$$
$$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

Goal

Estimate $\mathcal{P}$ with a low-rank approximation of $\mathcal{T}$.

Problem

What does “low rank” mean for a tensor?

Spiked Random Tensor Model

$$\mathcal{T} = \underbrace{\mathcal{P}}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$$N = n_1 + n_2 + n_3$$

$$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

Goal

Estimate $\mathcal{P}$ with a low-rank approximation of $\mathcal{T}$.

Problem

What does “low rank” mean for a tensor?

- Matrix case: rank-r

$$\text{SVD} \quad \rightsquigarrow \quad \boldsymbol{P} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top$$

Spiked Random Tensor Model

$$\mathcal{T} = \underbrace{\mathcal{P}}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$N = n_1 + n_2 + n_3$

$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$

Goal

Estimate $\mathcal{P}$ with a low-rank approximation of $\mathcal{T}$.

Problem

What does “low rank” mean for a tensor?

- Matrix case: rank-r

$$\text{SVD} \quad \rightsquigarrow \quad \boldsymbol{P} = \boldsymbol{U}\overset{r \times r}{\Sigma}\boldsymbol{V}^\top$$

Spiked Random Tensor Model

$$\mathcal{T} = \underbrace{\mathcal{P}}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$$N = n_1 + n_2 + n_3$$

$$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

Goal

Estimate $\mathcal{P}$ with a low-rank approximation of $\mathcal{T}$.

Problem

What does “low rank” mean for a tensor?

- Matrix case: rank-r

$$\text{SVD} \quad \rightsquigarrow \quad \boldsymbol{P} = \boldsymbol{U} \overset{r \times r}{\Sigma} \boldsymbol{V}^\top = [\![\Sigma; \boldsymbol{U}, \boldsymbol{V}]\!]$$

Spiked Random Tensor Model

$$\mathcal{T} = \underbrace{\mathcal{P}}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$$N = n_1 + n_2 + n_3$$

$$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

Goal

Estimate $\mathcal{P}$ with a low-rank approximation of $\mathcal{T}$.

Problem

What does "low rank" mean for a tensor?

- Matrix case: rank-r

$r \times r$

$$\text{SVD} \quad \rightsquigarrow \quad \boldsymbol{P} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top = [\![\boldsymbol{\Sigma}; \boldsymbol{U}, \boldsymbol{V}]\!]$$

- Tensor case: rank-(r_1, r_2, r_3)

$$\text{HOSVD} \quad \rightsquigarrow \quad \mathcal{P} = [\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!]$$

Spiked Random Tensor Model

$$\mathcal{T} = \underbrace{\mathcal{P}}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$$N = n_1 + n_2 + n_3$$
$$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

Goal

Estimate $\mathcal{P}$ with a low-rank approximation of $\mathcal{T}$.

Problem

What does "low rank" mean for a tensor?

- Matrix case: rank-r

↱ $r \times r$

$$\text{SVD} \quad \rightsquigarrow \quad \boldsymbol{P} = \boldsymbol{U}\Sigma\boldsymbol{V}^\top = [\![\Sigma; \boldsymbol{U}, \boldsymbol{V}]\!]$$

- Tensor case: rank-(r_1, r_2, r_3)

↱ Core tensor $r_1 \times r_2 \times r_3$

$$\text{HOSVD} \quad \rightsquigarrow \quad \mathcal{P} = [\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!]$$

Spiked Random Tensor Model

$$\mathcal{T} = \underbrace{\mathcal{P}}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$N = n_1 + n_2 + n_3$

$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$

Goal

Estimate $\mathcal{P}$ with a low-rank approximation of $\mathcal{T}$.

Problem

What does "low rank" mean for a tensor?

- Matrix case: rank-r

$\rightarrow r \times r$

$$\text{SVD} \quad \rightsquigarrow \quad \boldsymbol{P} = \boldsymbol{U}\boxed{\Sigma}\boldsymbol{V}^\top = [\![\Sigma; \boldsymbol{U}, \boldsymbol{V}]\!]$$

- Tensor case: rank-(r_1, r_2, r_3)

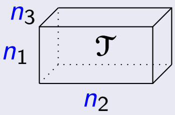
$\rightarrow$ Core tensor $r_1 \times r_2 \times r_3$

$$\text{HOSVD} \quad \rightsquigarrow \quad \mathcal{P} = [\![\boxed{\mathcal{H}}; \underbrace{\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}}]\!]$$

$\rightarrow$ Semi-orthogonal matrices $n_\ell \times r_\ell$

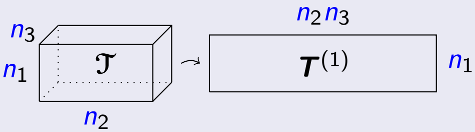
Higher-Order Singular Value Decomposition (HOSVD)

Tensor Unfolding



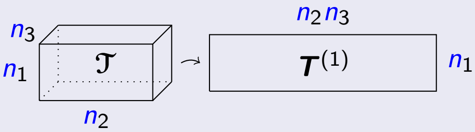
Higher-Order Singular Value Decomposition (HOSVD)

Tensor Unfolding



Higher-Order Singular Value Decomposition (HOSVD)

Tensor Unfolding

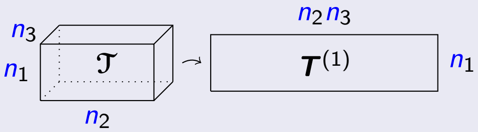


HOSVD

- $\boldsymbol{U} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(1)}$
- $\boldsymbol{V} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(2)}$
- $\boldsymbol{W} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(3)}$

Higher-Order Singular Value Decomposition (HOSVD)

Tensor Unfolding

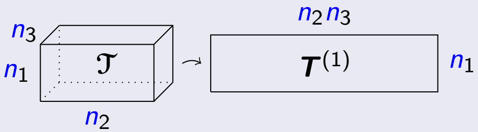


HOSVD

- $\boldsymbol{U} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(1)}$
- $\boldsymbol{V} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(2)}$
- $\boldsymbol{W} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(3)}$
- $\mathcal{G}_{a,b,c} \leftarrow \sum_{i,j,k=1}^{n_1,n_2,n_3} \mathcal{T}_{i,j,k} \boldsymbol{U}_{i,a} \boldsymbol{V}_{j,b} \boldsymbol{W}_{k,c}$

Higher-Order Singular Value Decomposition (HOSVD)

Tensor Unfolding

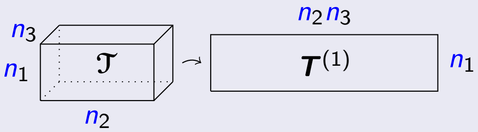


HOSVD

- $\boldsymbol{U} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(1)}$
- $\boldsymbol{V} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(2)}$
- $\boldsymbol{W} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(3)}$
- $\mathcal{G}_{a,b,c} \leftarrow \sum_{i,j,k=1}^{n_1,n_2,n_3} \mathcal{T}_{i,j,k} \boldsymbol{U}_{i,a} \boldsymbol{V}_{j,b} \boldsymbol{W}_{k,c}$
- $\mathcal{T} = [\![\mathcal{G}; \boldsymbol{U}, \boldsymbol{V}, \boldsymbol{W}]\!]$

Higher-Order Singular Value Decomposition (HOSVD)

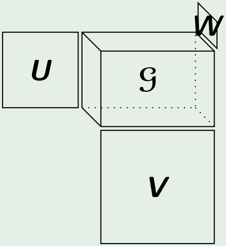
Tensor Unfolding



HOSVD

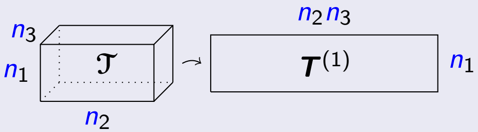
- $\boldsymbol{U} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(1)}$
- $\boldsymbol{V} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(2)}$
- $\boldsymbol{W} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(3)}$
- $\mathcal{G}_{a,b,c} \leftarrow \sum_{i,j,k=1}^{n_1,n_2,n_3} \mathcal{T}_{i,j,k} \boldsymbol{U}_{i,a} \boldsymbol{V}_{j,b} \boldsymbol{W}_{k,c}$
- $\mathcal{T} = [\![\mathcal{G}; \boldsymbol{U}, \boldsymbol{V}, \boldsymbol{W}]\!]$

HOSVD



Higher-Order Singular Value Decomposition (HOSVD)

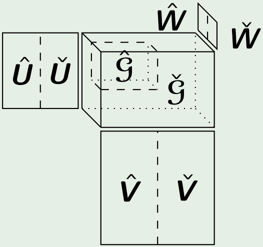
Tensor Unfolding



HOSVD

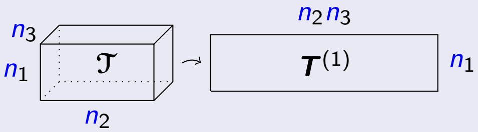
- $\boldsymbol{U} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(1)}$
- $\boldsymbol{V} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(2)}$
- $\boldsymbol{W} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(3)}$
- $\mathcal{G}_{a,b,c} \leftarrow \sum_{i,j,k=1}^{n_1,n_2,n_3} \mathcal{T}_{i,j,k} \boldsymbol{U}_{i,a} \boldsymbol{V}_{j,b} \boldsymbol{W}_{k,c}$
- $\mathcal{T} = [\![\mathcal{G}; \boldsymbol{U}, \boldsymbol{V}, \boldsymbol{W}]\!]$

Truncated HOSVD



Higher-Order Singular Value Decomposition (HOSVD)

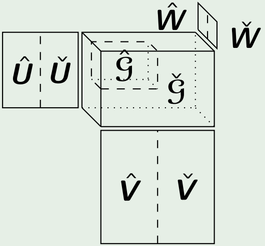
Tensor Unfolding



HOSVD

- $\boldsymbol{U} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(1)}$
- $\boldsymbol{V} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(2)}$
- $\boldsymbol{W} \leftarrow$ left singular vectors of $\boldsymbol{T}^{(3)}$
- $\mathcal{G}_{a,b,c} \leftarrow \sum_{i,j,k=1}^{n_1,n_2,n_3} \mathcal{T}_{i,j,k} \boldsymbol{U}_{i,a} \boldsymbol{V}_{j,b} \boldsymbol{W}_{k,c}$
- $\mathcal{T} = [\![\mathcal{G}; \boldsymbol{U}, \boldsymbol{V}, \boldsymbol{W}]\!]$

Truncated HOSVD



Multilinear Rank

$\hat{\mathcal{T}} = \left[\!\!\left[\hat{\mathcal{G}}; \hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}\right]\!\!\right]$ has *multilinear rank* (r_1, r_2, r_3).

Quasi-Optimality of the Truncated HOSVD

Proposition

$\hat{\mathfrak{T}} = \left[\!\!\left[\hat{\mathfrak{G}}; \hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}} \right]\!\!\right]$ is *not* the best low-rank approximation of $\mathfrak{T}$.

Quasi-Optimality of the Truncated HOSVD

Proposition

$\hat{\mathcal{T}} = \left[\!\!\left[\hat{\mathcal{G}}; \hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}\right]\!\!\right]$ is *not* the best low-rank approximation of $\mathcal{T}$.

Let $\mathcal{T}_\star$ be solution to the *maximum likelihood estimation* problem

$$\min_{\operatorname{rank}\mathcal{X} \leqslant (r_1, r_2, r_3)} \|\mathcal{T} - \mathcal{X}\|_{\mathsf{F}} \, .$$

Quasi-Optimality of the Truncated HOSVD

Proposition

$\hat{\mathcal{T}} = \left\llbracket \hat{\mathcal{G}}; \hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}} \right\rrbracket$ is *not* the best low-rank approximation of $\mathcal{T}$.

Let $\mathcal{T}_\star$ be solution to the *maximum likelihood estimation* problem

$$\min_{\operatorname{rank} \mathcal{X} \leqslant (r_1, r_2, r_3)} \|\mathcal{T} - \mathcal{X}\|_{\mathsf{F}} .$$

	MLE $\mathcal{T}_\star$	Truncated HOSVD $\hat{\mathcal{T}}$
Computation	✗ NP-Hard	✓ Polynomial time

Quasi-Optimality of the Truncated HOSVD

Proposition

$\hat{\mathcal{T}} = \left[\!\left[\hat{\mathcal{G}}; \hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}\right]\!\right]$ is *not* the best low-rank approximation of $\mathcal{T}$.

Let $\mathcal{T}_\star$ be solution to the *maximum likelihood estimation* problem

$$\min_{\operatorname{rank}\mathcal{X} \leqslant (r_1, r_2, r_3)} \|\mathcal{T} - \mathcal{X}\|_{\mathsf{F}}\,.$$

	MLE $\mathcal{T}_\star$	Truncated HOSVD $\hat{\mathcal{T}}$
Computation	✗ NP-Hard	✓ Polynomial time
Optimality	✓ Optimal	🆗 Quasi-optimal

Quasi-Optimality of the Truncated HOSVD

Proposition

$\hat{\mathcal{T}} = \left[\!\!\left[\hat{\mathcal{G}}; \hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}} \right]\!\!\right]$ is *not* the best low-rank approximation of $\mathcal{T}$.

Let $\mathcal{T}_\star$ be solution to the *maximum likelihood estimation* problem

$$\min_{\operatorname{rank} \mathcal{X} \leqslant (r_1, r_2, r_3)} \|\mathcal{T} - \mathcal{X}\|_{\mathsf{F}} \, .$$

	MLE $\mathcal{T}_\star$	Truncated HOSVD $\hat{\mathcal{T}}$
Computation	✗ NP-Hard	✅ Polynomial time
Optimality	✅ Optimal	🆗 Quasi-optimal

$$\|\mathcal{T} - \mathcal{T}_\star\|_{\mathsf{F}} \leqslant \left\|\mathcal{T} - \hat{\mathcal{T}}\right\|_{\mathsf{F}} \leqslant \sqrt{d}\, \|\mathcal{T} - \mathcal{T}_\star\|_{\mathsf{F}}$$

Spiked Tensor Model Analysis

$$\mathcal{T} = \underbrace{\mathcal{P}}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$$N = n_1 + n_2 + n_3$$

$$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

Spiked Tensor Model Analysis

$$\mathcal{T} = \underbrace{[\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!]}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$$N = n_1 + n_2 + n_3$$

$$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

Spiked Tensor Model Analysis

$$\mathcal{T} = \underbrace{[\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!]}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$$N = n_1 + n_2 + n_3$$

$$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$$

Question

Given $\hat{\mathcal{T}} = \left[\!\!\left[\hat{\mathcal{G}}; \hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}\right]\!\!\right]$, how close are $\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}$ from $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$?

Spiked Tensor Model Analysis

$$\mathcal{T} = \underbrace{[\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!]}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$$N = n_1 + n_2 + n_3$$
$$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

Question

Given $\hat{\mathcal{T}} = \left[\!\!\left[\hat{\mathcal{G}}; \hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}\right]\!\!\right]$, how close are $\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}$ from $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$?

Assumptions

- **Asymptotic regime**: $n_\ell \to +\infty$ with $0 < c_\ell = \lim n_\ell / N < +\infty$.

Spiked Tensor Model Analysis

$$\mathcal{T} = \underbrace{[\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!]}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$$N = n_1 + n_2 + n_3$$

$$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

Question

Given $\hat{\mathcal{T}} = \left[\!\!\left[\hat{\mathcal{G}}; \hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}\right]\!\!\right]$, how close are $\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}$ from $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$?

Assumptions

- **Asymptotic regime**: $n_\ell \to +\infty$ with $0 < c_\ell = \lim n_\ell / N < +\infty$.
- **Low-rank approximation**: $r_\ell \ll n_\ell$.

Spiked Tensor Model Analysis

$$\mathcal{T} = \underbrace{[\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!]}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3}, \qquad \begin{array}{l} N = n_1 + n_2 + n_3 \\ \mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1) \end{array}$$

Question

Given $\hat{\mathcal{T}} = \left[\!\!\left[\hat{\mathcal{G}}; \hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}\right]\!\!\right]$, how close are $\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}$ from $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$?

Assumptions

- **Asymptotic regime**: $n_\ell \to +\infty$ with $0 < c_\ell = \lim n_\ell / N < +\infty$.
- **Low-rank approximation**: $r_\ell \ll n_\ell$.
- **Non-triviality condition**: $\|\mathcal{P}\|_\mathsf{F} = \mathcal{O}(N^{\frac{d-2}{4}}) \quad (\mathbb{E}\left\|\frac{1}{\sqrt{N}}\mathcal{N}\right\|_\mathsf{F} = \mathcal{O}(N^{\frac{d-1}{2}}))$.

Limiting Spectral Distribution

- $\mathcal{T} = [\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!] + \frac{1}{\sqrt{N}}\mathcal{N}.$

Limiting Spectral Distribution

- $\mathcal{T} = [\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!] + \frac{1}{\sqrt{N}}\mathcal{N}$.
- $\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}$ gather the first left singular vectors of $\boldsymbol{T}^{(1)}, \boldsymbol{T}^{(2)}, \boldsymbol{T}^{(3)}$.

Limiting Spectral Distribution

- $\mathcal{T} = [\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!] + \frac{1}{\sqrt{N}}\mathcal{N}$.
- $\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}$ gather the first left singular vectors of $\boldsymbol{T}^{(1)}, \boldsymbol{T}^{(2)}, \boldsymbol{T}^{(3)}$.

➡ We study the limiting spectrum of $\boldsymbol{T}^{(\ell)} {\boldsymbol{T}^{(\ell)}}^{\top}$.

Limiting Spectral Distribution

- $\mathcal{T} = [\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!] + \frac{1}{\sqrt{N}}\mathcal{N}$.
- $\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}$ gather the first left singular vectors of $\boldsymbol{T}^{(1)}, \boldsymbol{T}^{(2)}, \boldsymbol{T}^{(3)}$.

➡ We study the limiting spectrum of $\boldsymbol{T}^{(\ell)} \boldsymbol{T}^{(\ell)\top}$.

⚠ The dimensions of $\boldsymbol{T}^{(\ell)}$ do *not* have similar sizes: $N \ll N^{d-1}$!

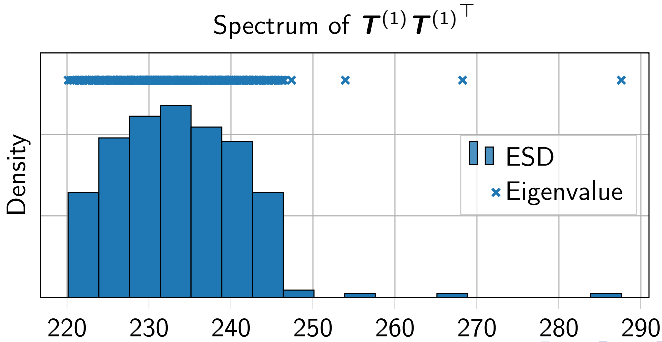
Limiting Spectral Distribution

- $\mathcal{T} = [\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!] + \frac{1}{\sqrt{N}}\mathcal{N}$.
- $\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}$ gather the first left singular vectors of $\boldsymbol{T}^{(1)}, \boldsymbol{T}^{(2)}, \boldsymbol{T}^{(3)}$.

➡ We study the limiting spectrum of $\boldsymbol{T}^{(\ell)}\boldsymbol{T}^{(\ell)\top}$.

⚠ The dimensions of $\boldsymbol{T}^{(\ell)}$ do *not* have similar sizes: $N \ll N^{d-1}$!

- Experiment with $(n_1, n_2, n_3) = (300, 500, 700)$, $(r_1, r_2, r_3) = (3, 4, 5)$.



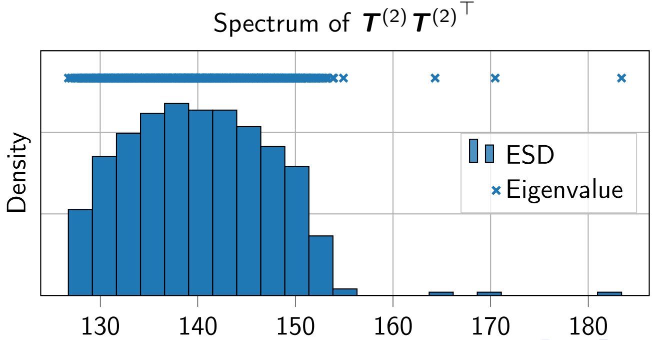
Limiting Spectral Distribution

- $\mathcal{T} = [\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!] + \frac{1}{\sqrt{N}}\mathcal{N}$.
- $\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}$ gather the first left singular vectors of $\boldsymbol{T}^{(1)}, \boldsymbol{T}^{(2)}, \boldsymbol{T}^{(3)}$.

➡ We study the limiting spectrum of $\boldsymbol{T}^{(\ell)}\boldsymbol{T}^{(\ell)\top}$.

⚠ The dimensions of $\boldsymbol{T}^{(\ell)}$ do *not* have similar sizes: $N \ll N^{d-1}$!

- Experiment with $(n_1, n_2, n_3) = (300, 500, 700)$, $(r_1, r_2, r_3) = (3, 4, 5)$.



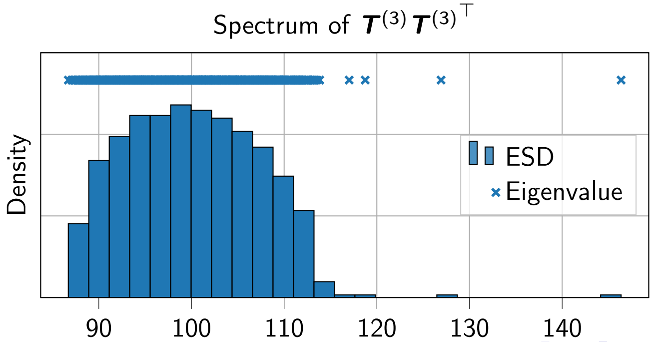
Limiting Spectral Distribution

- $\mathcal{T} = [\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!] + \frac{1}{\sqrt{N}}\mathcal{N}$.
- $\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}$ gather the first left singular vectors of $\boldsymbol{T}^{(1)}, \boldsymbol{T}^{(2)}, \boldsymbol{T}^{(3)}$.

➡ We study the limiting spectrum of $\boldsymbol{T}^{(\ell)}\boldsymbol{T}^{(\ell)\top}$.

⚠ The dimensions of $\boldsymbol{T}^{(\ell)}$ do *not* have similar sizes: $N \ll N^{d-1}$!

- Experiment with $(n_1, n_2, n_3) = (300, 500, 700)$, $(r_1, r_2, r_3) = (3, 4, 5)$.



Limiting Spectral Distribution

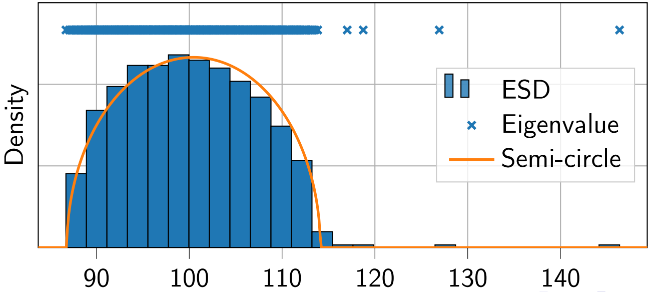
- $\mathcal{T} = [\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!] + \frac{1}{\sqrt{N}}\mathcal{N}$.
- $\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}$ gather the first left singular vectors of $\boldsymbol{T}^{(1)}, \boldsymbol{T}^{(2)}, \boldsymbol{T}^{(3)}$.

➡ We study the limiting spectrum of $\boldsymbol{T}^{(\ell)}\boldsymbol{T}^{(\ell)\top}$.

⚠ The dimensions of $\boldsymbol{T}^{(\ell)}$ do *not* have similar sizes: $N \ll N^{d-1}$!

- Experiment with $(n_1, n_2, n_3) = (300, 500, 700)$, $(r_1, r_2, r_3) = (3, 4, 5)$.

Spectrum of $\boldsymbol{T}^{(3)}\boldsymbol{T}^{(3)\top}$



Limiting Spectral Distribution

- $\mathcal{T} = [\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!] + \frac{1}{\sqrt{N}}\mathcal{N}$.
- $\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}$ gather the first left singular vectors of $\boldsymbol{T}^{(1)}, \boldsymbol{T}^{(2)}, \boldsymbol{T}^{(3)}$.

➡ We study the limiting spectrum of $\boldsymbol{T}^{(\ell)}\boldsymbol{T}^{(\ell)\top}$.

⚠ The dimensions of $\boldsymbol{T}^{(\ell)}$ do *not* have similar sizes: $N \ll N^{d-1}$!

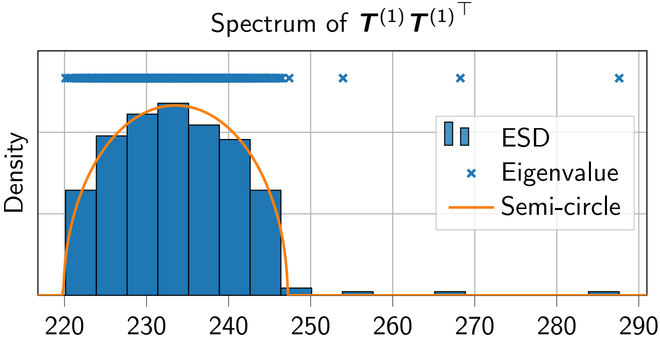
Theorem (Limiting Spectral Distribution)

$$\mu_N^{(\ell)} = \frac{n_1 n_2 n_3}{n_\ell N} = \mathcal{O}(N^{d-2}), \qquad \sigma_N = \frac{\sqrt{n_1 n_2 n_3}}{N} = \mathcal{O}(N^{\frac{d-2}{2}}).$$

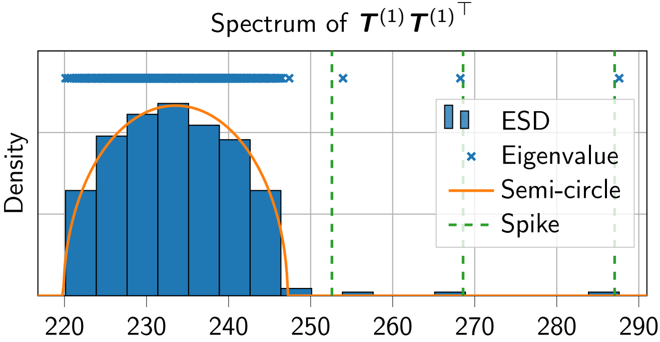
As $N \to +\infty$, $\frac{1}{\sigma_N}\boldsymbol{T}^{(\ell)}\boldsymbol{T}^{(\ell)\top} - \frac{\mu_N^{(\ell)}}{\sigma_N}\boldsymbol{I}_{n_\ell}$ has a limiting spectral distribution $\tilde{\nu}$ whose Stieltjes transform $\tilde{m}$ satisfies

$$\tilde{m}^2(\tilde{z}) + \tilde{z}\tilde{m}(\tilde{z}) + 1 = 0 \qquad \tilde{z} \in \mathbb{C} \setminus \operatorname{supp} \tilde{\nu}$$

Spikes Behavior



Spikes Behavior

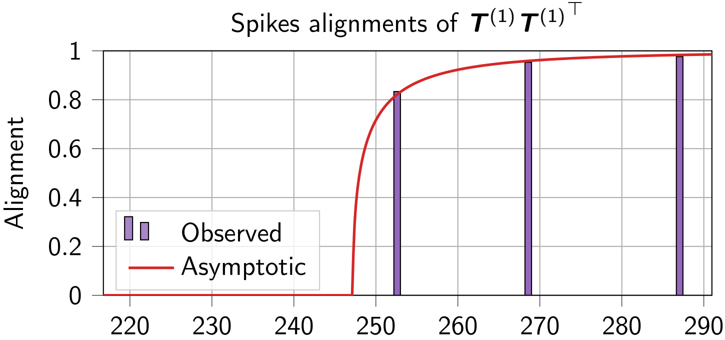


Theorem (Spikes Behavior)

If $\rho_{\kappa,N}^{(\ell)} \equiv \frac{s_\kappa^2(\boldsymbol{P}^{(\ell)})}{\sigma_N} > 1$, *then*

$$s_\kappa^2(\boldsymbol{T}^{(\ell)}) \underset{N\to+\infty}{\overset{a.s.}{\sim}} \rho_{\kappa,N}^{(\ell)} + \frac{1}{\rho_{\kappa,N}^{(\ell)}}, \qquad \left\langle \hat{\boldsymbol{u}}_{\kappa,N}^{(\ell)}, \boldsymbol{x}_\kappa^{(\ell)} \right\rangle^2 \underset{N\to+\infty}{\overset{a.s.}{\sim}} 1 - \frac{1}{\rho_{\kappa,N}^{(\ell)\,2}}.$$

Spikes Behavior

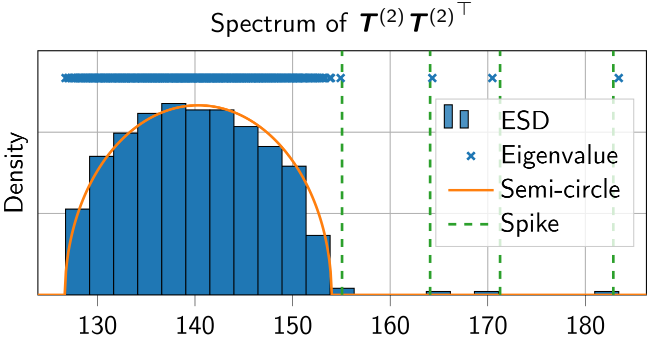


Theorem (Spikes Behavior)

If $\rho_{\kappa,N}^{(\ell)} \equiv \frac{s_\kappa^2(\boldsymbol{P}^{(\ell)})}{\sigma_N} > 1$, *then*

$$s_\kappa^2(\boldsymbol{T}^{(\ell)}) \underset{N\to+\infty}{\overset{a.s.}{\sim}} \rho_{\kappa,N}^{(\ell)} + \frac{1}{\rho_{\kappa,N}^{(\ell)}}, \qquad \left\langle \hat{\boldsymbol{u}}_{\kappa,N}^{(\ell)}, \boldsymbol{x}_\kappa^{(\ell)} \right\rangle^2 \underset{N\to+\infty}{\overset{a.s.}{\sim}} 1 - \frac{1}{{\rho_{\kappa,N}^{(\ell)}}^2}.$$

Spikes Behavior

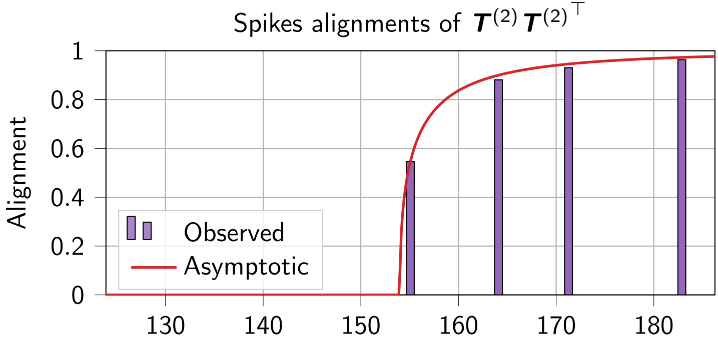


Theorem (Spikes Behavior)

If $\rho_{\kappa,N}^{(\ell)} \equiv \frac{s_\kappa^2(\boldsymbol{P}^{(\ell)})}{\sigma_N} > 1$*, then*

$$s_\kappa^2(\boldsymbol{T}^{(\ell)}) \underset{N\to+\infty}{\overset{a.s.}{\sim}} \rho_{\kappa,N}^{(\ell)} + \frac{1}{\rho_{\kappa,N}^{(\ell)}}, \qquad \left\langle \hat{\boldsymbol{u}}_{\kappa,N}^{(\ell)}, \boldsymbol{x}_\kappa^{(\ell)} \right\rangle^2 \underset{N\to+\infty}{\overset{a.s.}{\sim}} 1 - \frac{1}{\rho_{\kappa,N}^{(\ell)\;2}}.$$

Spikes Behavior

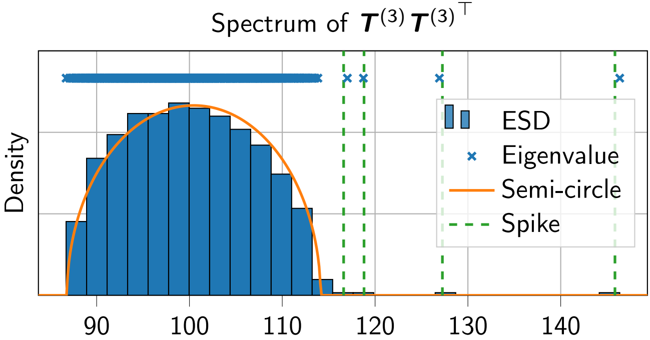

Spikes alignments of $\boldsymbol{T}^{(2)}\boldsymbol{T}^{(2)^\top}$

Theorem (Spikes Behavior)

If $\rho_{\kappa,N}^{(\ell)} \equiv \frac{s_\kappa^2(\boldsymbol{P}^{(\ell)})}{\sigma_N} > 1,$ *then*

$$s_\kappa^2(\boldsymbol{T}^{(\ell)}) \underset{N\to+\infty}{\overset{a.s.}{\sim}} \rho_{\kappa,N}^{(\ell)} + \frac{1}{\rho_{\kappa,N}^{(\ell)}}, \qquad \left\langle \hat{\boldsymbol{u}}_{\kappa,N}^{(\ell)}, \boldsymbol{x}_\kappa^{(\ell)} \right\rangle^2 \underset{N\to+\infty}{\overset{a.s.}{\sim}} 1 - \frac{1}{\rho_{\kappa,N}^{(\ell)}{}^2}.$$

Spikes Behavior

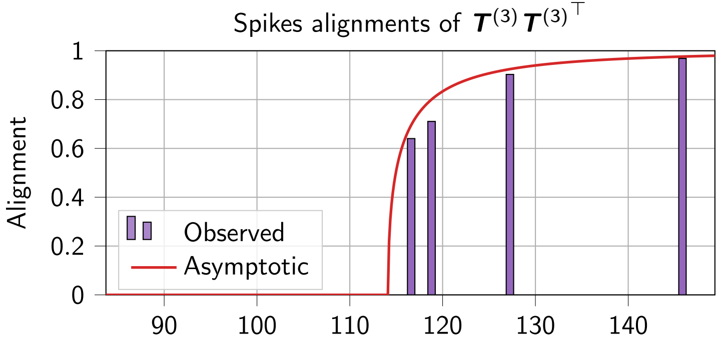


Theorem (Spikes Behavior)

If $\rho^{(\ell)}_{\kappa,N} \equiv \frac{s^2_\kappa(\boldsymbol{P}^{(\ell)})}{\sigma_N} > 1$, *then*

$$s^2_\kappa(\boldsymbol{T}^{(\ell)}) \underset{N\to+\infty}{\overset{a.s.}{\sim}} \rho^{(\ell)}_{\kappa,N} + \frac{1}{\rho^{(\ell)}_{\kappa,N}}, \qquad \left\langle \hat{\boldsymbol{u}}^{(\ell)}_{\kappa,N}, \boldsymbol{x}^{(\ell)}_\kappa \right\rangle^2 \underset{N\to+\infty}{\overset{a.s.}{\sim}} 1 - \frac{1}{{\rho^{(\ell)}_{\kappa,N}}^2}.$$

Spikes Behavior



Theorem (Spikes Behavior)

If $\rho^{(\ell)}_{\kappa,N} \equiv \frac{s^2_\kappa(\boldsymbol{P}^{(\ell)})}{\sigma_N} > 1$, *then*

$$s^2_\kappa(\boldsymbol{T}^{(\ell)}) \underset{N\to+\infty}{\overset{a.s.}{\sim}} \rho^{(\ell)}_{\kappa,N} + \frac{1}{\rho^{(\ell)}_{\kappa,N}}, \qquad \left\langle \hat{\boldsymbol{u}}^{(\ell)}_{\kappa,N}, \boldsymbol{x}^{(\ell)}_\kappa \right\rangle^2 \underset{N\to+\infty}{\overset{a.s.}{\sim}} 1 - \frac{1}{\rho^{(\ell)\,2}_{\kappa,N}}.$$

Conclusion

$$\mathcal{T} = \underbrace{[\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!]}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$$N = n_1 + n_2 + n_3$$

$$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$$

Question

Given $\hat{\mathcal{T}} = \left[\!\!\left[\hat{\mathcal{G}}; \hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}\right]\!\!\right]$, how close are $\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}$ from $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$?

Conclusion

$$\mathcal{T} = \underbrace{[\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!]}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$$N = n_1 + n_2 + n_3$$

$$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$$

Question

Given $\hat{\mathcal{T}} = \left[\!\!\left[\hat{\mathcal{G}}; \hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}\right]\!\!\right]$, how close are $\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}$ from $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$?

Answer — Subspace Alignments

$$\frac{1}{r_1}\left\|\boldsymbol{X}^\top \hat{\boldsymbol{U}}\right\|_{\mathrm{F}}^2$$

Conclusion

$$\mathcal{T} = \underbrace{[\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!]}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$$N = n_1 + n_2 + n_3$$
$$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$$

Question

Given $\hat{\mathcal{T}} = \left[\!\!\left[\hat{\mathcal{G}}; \hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}\right]\!\!\right]$, how close are $\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}$ from $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$?

Answer — Subspace Alignments

$$\frac{1}{r_1}\left\|\boldsymbol{X}^\top \hat{\boldsymbol{U}}\right\|_{\mathrm{F}}^2 = \frac{1}{r_1}\sum_{\kappa=1}^{r_1} \cos^2 \theta_{\kappa, N}^{(1)}(\boldsymbol{X}, \hat{\boldsymbol{U}})$$

Conclusion

$$\mathcal{T} = \underbrace{[\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!]}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$$N = n_1 + n_2 + n_3$$

$$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

Question

Given $\hat{\mathcal{T}} = \left[\!\!\left[\hat{\mathcal{G}}; \hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}\right]\!\!\right]$, how close are $\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}$ from $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$?

Answer — Subspace Alignments

$$\frac{1}{r_1}\left\|\boldsymbol{X}^\top \hat{\boldsymbol{U}}\right\|_{\mathrm{F}}^2 = \frac{1}{r_1}\sum_{\kappa=1}^{r_1} \cos^2 \boxed{\theta^{(1)}_{\kappa,N}(\boldsymbol{X}, \hat{\boldsymbol{U}})}$$

κ-th principal angle

Conclusion

$$\mathcal{T} = \underbrace{[\![\mathcal{H}; \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}]\!]}_{\text{signal (low rank)}} + \underbrace{\frac{1}{\sqrt{N}}\mathcal{N}}_{\text{noise (random)}} \in \mathbb{R}^{n_1 \times n_2 \times n_3},$$

$$N = n_1 + n_2 + n_3$$
$$\mathcal{N}_{i,j,k} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$$

Question

Given $\hat{\mathcal{T}} = \left[\!\!\left[\hat{\mathcal{G}}; \hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}\right]\!\!\right]$, how close are $\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}$ from $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$?

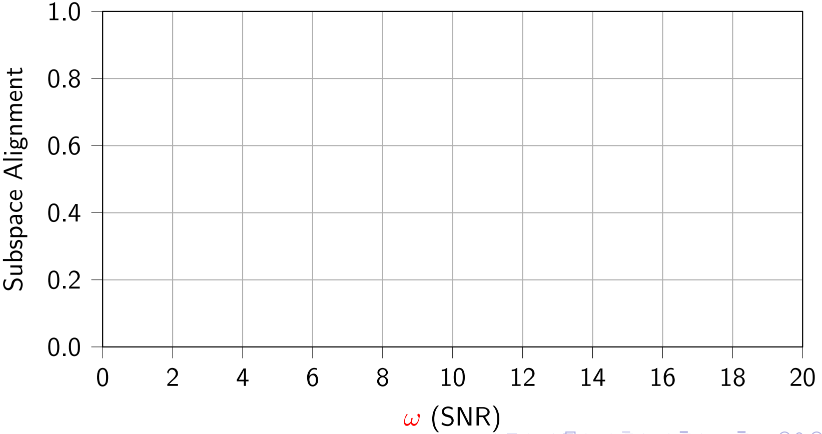
Answer — Subspace Alignments

$$\frac{1}{r_1}\left\|\boldsymbol{X}^\top \hat{\boldsymbol{U}}\right\|_{\mathrm{F}}^2 = \frac{1}{r_1}\sum_{\kappa=1}^{r_1} \cos^2 \boxed{\theta_{\kappa,N}^{(1)}(\boldsymbol{X}, \hat{\boldsymbol{U}})} \underset{N \to +\infty}{\overset{\text{a.s.}}{\sim}} \frac{1}{r_1}\sum_{\kappa=1}^{r_1}\left[1 - \frac{1}{\rho_{\kappa,N}^{(1)\,2}}\right]^+$$

κ-th principal angle

Perspectives

$$\mathfrak{T} = \sqrt{\omega}\mathcal{P}_0 + \frac{1}{\sqrt{N}}\mathcal{N} \qquad \text{with} \qquad \|\mathcal{P}_0\|_{\mathsf{F}}^2 = \sigma_N$$

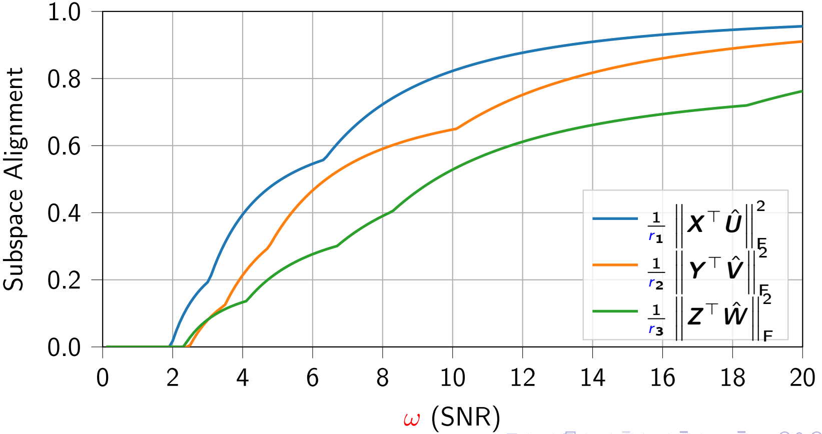


Perspectives

$$\mathcal{T} = \sqrt{\omega}\mathcal{P}_0 + \frac{1}{\sqrt{N}}\mathcal{N} \qquad \text{with} \qquad \|\mathcal{P}_0\|_\mathsf{F}^2 = \sigma_N$$

Truncated HOSVD

$$\hat{\mathcal{T}} = \left[\!\left[\hat{\mathcal{G}}; \hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}\right]\!\right]$$



Perspectives

$$\mathfrak{T} = \sqrt{\omega}\,\mathcal{P}_0 + \frac{1}{\sqrt{N}}\mathcal{N} \qquad \text{with} \qquad \|\mathcal{P}_0\|_{\mathrm{F}}^2 = \sigma_N$$

Truncated HOSVD

$$\hat{\mathfrak{T}} = [\![\hat{\mathcal{G}}; \hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}, \hat{\boldsymbol{W}}]\!]$$

Maximum Likelihood

$$\mathfrak{T}_\star = [\![\mathcal{G}_\star; \boldsymbol{U}_\star, \boldsymbol{V}_\star, \boldsymbol{W}_\star]\!]$$