Robust Tensor Decomposition via Orientation Invariant Tubal Nuclear Norms

WANG Andong,\textsuperscript{1} LI Chao,\textsuperscript{2} JIN Zhong,\textsuperscript{1} ZHAO Qibin\textsuperscript{2}

\textsuperscript{1}Nanjing University of Science and Technology, China
\textsuperscript{2}Tensor Learning Unit, RIKEN AIP, Japan
1 Motivation
   - Robust Tensor Decomposition
   - Low-tubal-rank Structure

2 Orientation Invariant TNNs for RTD
   - Orientation Invariant TNNs
   - Proposed RTD Models
   - Error bounds

3 Experiments

4 Conclusion
Tensor data is everywhere!

Motivation

Robust Tensor Decomposition

Psychology
Behavior analysis

Process monitoring
Failure detection

Environment monitoring
Quality assessment

Video surveillance
Anomaly detection

Image/Video processing
Inpainting/De-noising

Social networks
Community detection

Question system
Topic model

EEG signal processing
Disease surveillance

MRI
Behavior recognition

(NJU&RIKEN TLU)  OITNN for RTD
Robust Tensor Decomposition (RTD)

⚠️ Observed tensor data are often not clean
May be corrupted by both outliers and noises
Due to: sensor failures, lens pollution, video abnormalities, corruption of images, ...

😊 Many tensor data are low-rank
E.g. images and videos have (well/approx.) low-rank structure
(Liu J et al. PAMI 2013; Zhao QB et al. PAMI 2015)

↓ This paper

An Observation Model (Gu QQ et al. NIPS 2014)

\[ \mathcal{Y} = \mathcal{L}^* + \mathcal{S}^* + \mathcal{E} \in \mathbb{R}^{d_1 \times \cdots \times d_K} \]
Robust Tensor Decomposition (RTD)

Motivation

Robust Tensor Decomposition

Observation Model

\[ Y = L^* + S^* + E \in \mathbb{R}^{d_1 \times \cdots \times d_K} \]

Observed tensor data are often not clean

May be corrupted by both outliers and noises

Due to: sensor failures, lens pollution, video abnormalities, corruption of images, ...

Many tensor data are low-rank

E.g. images and videos have (well/approx.) low-rank structure

(Liu J et al. PAMI 2013; Zhao QB et al. PAMI 2015)

This paper

An Observation Model (Gu QQ et al. NIPS 2014)
Robust Tensor Decomposition (RTD)

Motivation

Robust Tensor Decomposition (RTD)

Robust Tensor Decomposition (RTD)

Observed tensor data are often not clean

May be corrupted by both outliers and noises

Due to: sensor failures, lens pollution, video abnormalities, corruption of images, ...

Many tensor data are low-rank

E.g. images and videos have (well/approx.) low-rank structure

(Liu J et al. PAMI 2013; Zhao QB et al. PAMI 2015)

↓ This paper

An Observation Model (Gu QQ et al. NIPS 2014)

\[ Y = \mathcal{L}^* + \mathcal{S}^* + \mathcal{E} \in \mathbb{R}^{d_1 \times \cdots \times d_K} \]
Robust Tensor Decomposition (RTD)

Motivation

Robust Tensor Decomposition

Observed tensor data are often not clean

May be corrupted by both outliers and noises

Due to: sensor failures, lens pollution, video abnormalities, corruption of images, ...

Many tensor data are low-rank

E.g. images and videos have (well/approx.) low-rank structure

(Liu J et al. PAMI 2013; Zhao QB et al. PAMI 2015)

An Observation Model (Gu QQ et al. NIPS 2014)

\[ \mathbf{Y} = \mathbf{L}^* + \mathbf{S}^* + \mathbf{E} \in \mathbb{R}^{d_1 \times \cdots \times d_K} \]
Motivation

Robust Tensor Decomposition (RTD)

🔥 Observed tensor data are often not clean

May be corrupted by both outliers and noises

Due to: sensor failures, lens pollution, video abnormalities, corruption of images, ...

😊 Many tensor data are low-rank

E.g. images and videos have (well/approx.) low-rank structure

(Liu J et al. PAMI 2013; Zhao QB et al. PAMI 2015)

↓ This paper

An Observation Model (Gu QQ et al. NIPS 2014)

\[ \mathcal{Y} = \mathcal{L}^* + \mathcal{S}^* + \mathcal{E} \in \mathbb{R}^{d_1 \times \cdots \times d_K} \]
Robust Tensor Decomposition (RTD)

Smiley face

Observed tensor data are often not clean

May be corrupted by both outliers and noises

Due to: sensor failures, lens pollution, video abnormalities, corruption of images, ...

Sad face

Many tensor data are low-rank

E.g. images and videos have (well/approx.) low-rank structure
(Liu J et al. PAMI 2013; Zhao QB et al. PAMI 2015)

This paper

An Observation Model (Gu QQ et al. NIPS 2014)

\[ \mathcal{Y} = \mathcal{L}^* + \mathcal{S}^* + \mathcal{E} \in \mathbb{R}^{d_1 \times \cdots \times d_K} \]
Robust Tensor Decomposition (RTD)

Motivation

Observed tensor data are often not clean

May be corrupted by both **outliers** and **noises**

Due to: sensor failures, lens pollution, video abnormalities, corruption of images, ...

Many tensor data are low-rank

E.g. images and videos have (well/approx.) **low-rank structure**

(Liu J et al. PAMI 2013; Zhao QB et al. PAMI 2015)

\[ \mathcal{Y} = \mathcal{L}^* + \mathcal{S}^* + \mathcal{E} \in \mathbb{R}^{d_1 \times \cdots \times d_K} \]

An Observation Model (Gu QQ et al. NIPS 2014)

- Observed tensor
- Low-rank tensor
- Sparse outliers
- Small noises

(NJUST&RIKEN TLU)
Robust Tensor Decomposition (RTD)

- Observed tensor data are often not clean
  May be corrupted by both **outliers** and **noises**
  Due to: sensor failures, lens pollution, video abnormalities, corruption of images, ...

- Many tensor data are low-rank
  E.g. images and videos have (well/approx.) **low-rank structure**
  \( \text{(Liu J et al. PAMI 2013; Zhao QB et al. PAMI 2015)} \)

\[ \mathbf{Y} = \mathbf{L}^* + \mathbf{S}^* + \mathbf{E} \in \mathbb{R}^{d_1 \times \ldots \times d_K} \]

\( \mathbf{Y} \)  \( \mathbf{L}^* \)  \( \mathbf{S}^* \)  \( \mathbf{E} \)
Robust Tensor Decomposition (RTD)

😊 Observed tensor data are often not clean

- May be corrupted by both **outliers** and **noises**
- Due to: sensor failures, lens pollution, video abnormalities, corruption of images, ...

😊 Many tensor data are low-rank

- E.g. images and videos have (well/approx.) **low-rank structure**
  (Liu J et al. PAMI 2013; Zhao QB et al. PAMI 2015)

↓ This paper

An Observation Model (Gu QQ et al. NIPS 2014)

\[ \mathcal{Y} = \mathcal{L}^* + \mathcal{S}^* + \mathcal{E} \in \mathbb{R}^{d_1 \times \cdots \times d_K} \]
Robust Tensor Decomposition

Motivation

Problem

How to estimate the clean $L^*$ from corrupted observation $Y \in \mathbb{R}^{d_1 \times \cdots \times d_K}$?

How to exploit the low-rank structure of $L^*$?
Robust Tensor Decomposition

**Motivation**

**Robust Tensor Decomposition**

**Problem**

How to estimate the clean $\mathcal{L}^*$ from corrupted observation $\mathcal{Y} \in \mathbb{R}^{d_1 \times \ldots \times d_K}$?

How to exploit the low-rank structure of $\mathcal{L}^*$?
Motivation

Low-tubal-rank Structure

Commonly used tensor low-rank structure

Motivation

Low-tubal-rank Structure

shown to have stronger modeling capabilities than low-Tucker-rank/low-CP-rank structure for images and videos\(^a\)


OITNN for RTD
Commonly used tensor low-rank structure

Motivation

Low-tubal-rank Structure

Low Tucker rank structure

Low CP rank structure

Low-tubal-rank Structure

shown to have stronger modeling capabilities than low-Tucker-rank/low-CP-rank structure for images and videos

Motivation

Low-tubal-rank Structure

Theorem 1 (Tensor SVD (Kilmer et al. 2013)).

Any 3-way tensor $\mathcal{T} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ can be decomposed as

$$\mathcal{T} = \mathcal{U} \ast \mathcal{S} \ast \mathcal{V}^\top$$

1. $\ast$ is the tensor-tensor product (t-product) (Kilmer et al. 2013)
2. $\mathcal{U} \in \mathbb{R}^{d_1 \times d_1 \times d_3}$, $\mathcal{V} \in \mathbb{R}^{d_2 \times d_2 \times d_3}$ are orthogonal tensors (Kilmer et al. 2013)
3. $\mathcal{S} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ is an $f$-diagonal tensor (Kilmer et al. 2013)
Any 3-way tensor $\mathcal{T} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ can be decomposed as

$$\mathcal{T} = \mathcal{U} \ast \mathcal{S} \ast \mathcal{V}^\top$$

1. $\ast$ is the tensor-tensor product (t-product) (Kilmer et al. 2013)
2. $\mathcal{U} \in \mathbb{R}^{d_1 \times d_1 \times d_3}$, $\mathcal{V} \in \mathbb{R}^{d_2 \times d_2 \times d_3}$ are orthogonal tensors (Kilmer et al. 2013)
3. $\mathcal{S} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ is an $f$-diagonal tensor (Kilmer et al. 2013)
Low-tubal-rank Structure

\[ \mathcal{T} = \mathcal{U} \ast \mathcal{S} \ast \mathcal{V}^\top \]

**Theorem 1 (Tensor SVD (Kilmer et al. 2013)).**

Any 3-way tensor \( \mathcal{T} \in \mathbb{R}^{d_1 \times d_2 \times d_3} \) can be decomposed as

\[ \mathcal{T} = \mathcal{U} \ast \mathcal{S} \ast \mathcal{V}^\top \]

1. \( \ast \) is the tensor-tensor product (t-product) (Kilmer et al. 2013)
2. \( \mathcal{U} \in \mathbb{R}^{d_1 \times d_1 \times d_3}, \mathcal{V} \in \mathbb{R}^{d_2 \times d_2 \times d_3} \) are orthogonal tensors (Kilmer et al. 2013)
3. \( \mathcal{S} \in \mathbb{R}^{d_1 \times d_2 \times d_3} \) is an \( f \)-diagonal tensor (Kilmer et al. 2013)
Motivation
Low-tubal-rank Structure

Low-tubal-rank Structure

\[ \mathbf{T} = \mathbf{U} \ast \mathbf{S} \ast \mathbf{V}^\top \]

**Theorem 1 (Tensor SVD (Kilmer et al. 2013)).**

Any 3-way tensor \( \mathbf{T} \in \mathbb{R}^{d_1 \times d_2 \times d_3} \) can be decomposed as

\[ \mathbf{T} = \mathbf{U} \ast \mathbf{S} \ast \mathbf{V}^\top \]

- \( \ast \) is the tensor-tensor product (t-product) (Kilmer et al. 2013)
- \( \mathbf{U} \in \mathbb{R}^{d_1 \times d_1 \times d_3} \), \( \mathbf{V} \in \mathbb{R}^{d_2 \times d_2 \times d_3} \) are orthogonal tensors (Kilmer et al. 2013)
- \( \mathbf{S} \in \mathbb{R}^{d_1 \times d_2 \times d_3} \) is an \( f \)-diagonal tensor (Kilmer et al. 2013)
**Motivation**

**Low-tubal-rank Structure**

Theorem 1 (Tensor SVD (Kilmer et al. 2013)).

Any 3-way tensor $\mathcal{T} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ can be decomposed as

$$\mathcal{T} = \mathcal{U} \ast S \ast \mathcal{V}^\top$$

1. $\ast$ is the tensor-tensor product (t-product) (Kilmer et al. 2013)
2. $\mathcal{U} \in \mathbb{R}^{d_1 \times d_1 \times d_3}$, $\mathcal{V} \in \mathbb{R}^{d_2 \times d_2 \times d_3}$ are orthogonal tensors (Kilmer et al. 2013)
3. $S \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ is an $f$-diagonal tensor (Kilmer et al. 2013)
Definition 2 (Tubal Rank (Kilmer et al. 2013)).
The tubal rank of $\mathcal{T} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ is the number of non-zero tubes in $\mathcal{S}$

$$r_{tb}(\mathcal{T}) := \#\{i \mid \mathcal{S}(i, i, :) \neq 0\}$$

Relationship between t-product and DFT indicates (Lu CY et al. PAMI 2019):

$$r_{tb}(\mathcal{T}) = \#\{i \mid \mathcal{S}(i, i, 1) \neq 0\}$$
Definition 2 (Tubal Rank (Kilmer et al. 2013)).

The tubal rank of $\mathcal{T} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ is the number of non-zero tubes in $\mathcal{S}$

$$r_{tb}(\mathcal{T}) := \# \{ i \mid \mathcal{S}(i, i, :) \neq 0 \}$$

Relationship between t-product and DFT indicates (Lu CY et al. PAMI 2019):

$$r_{tb}(\mathcal{T}) = \# \{ i \mid \mathcal{S}(i, i, 1) \neq 0 \}$$
Tensor “Singular Values”

\[ r_{tb}(\mathcal{T}) = \#\{i \mid S(i, i, 1) \neq 0\} \]

\( S(i, i, 1) \)'s are also called the "singular values" of tensor \( \mathcal{T} \) (Lu CY et al. PAMI 2019)

Definition 3 (Tubal Nuclear Norm, TNN).

The TNN of \( \mathcal{T} \) is the sum of its singular values

\[
\| \mathcal{T} \|_* := \sum_{i=1}^{d_1 \wedge d_2} S(i, i, 1)
\]
Tensor “Singular Values”

\[ r_{tb}(T) = \# \{ i \mid S(i, i, 1) \neq 0 \} \]

\( S(i, i, 1) \)'s are also called the “\textit{singular values}” of tensor \( T \) (Lu CY et al. PAMI 2019)

Definition 3 (Tubal Nuclear Norm, TNN).

The TNN of \( T \) is the sum of its singular values

\[ \| T \|_* := \sum_{i=1}^{d_1 \wedge d_2} S(i, i, 1) \]
Motivation

Low-tubal-rank Structure

Tensor “Singular Values”

\[ r_{tb}(\mathcal{T}) = \#\{i \mid S(i, i, 1) \neq 0\} \]

\( S(i, i, 1)'s \) are also called the “singular values” of tensor \( \mathcal{T} \) (Lu CY et al. PAMI 2019)

Definition 3 (Tubal Nuclear Norm, TNN).

The TNN of \( \mathcal{T} \) is the sum of its singular values

\[ \| \mathcal{T} \|_* := \sum_{i=1}^{d_1 \wedge d_2} S(i, i, 1) \]
Low-rankness in spectral domain

Relationship between $t$-product and DFT indicates:

$$\|T\|_* = \frac{1}{d_3} \sum_{k=1}^{d_3} \|\mathcal{T}(:, :, k)\|_*$$

TNN measures low-rankness in spectral domain along the 3d orientation
Low-rankness in spectral domain

Relationship between $t$-product and DFT indicates:

$$\|T\|_* = \frac{1}{d_3} \sum_{k=1}^{d_3} \|\tilde{T}(\cdot, \cdot, k)\|_*$$

TNN measures low-rankness in spectral domain along the 3d orientation.
Motivation

Low-tubal-rank Structure

Low-rankness in spectral domain

Relationship between $t$-product and DFT indicates:

$$\|T\|_* = \frac{1}{d_3} \sum_{k=1}^{d_3} \| \tilde{T}(\cdot, \cdot, k) \|_*$$

TNN measures low-rankness in spectral domain along the 3d orientation
Motivation

Low-tubal-rank Structure

Weaknesses of TNN

\[ \| \mathcal{T} \|_* = \frac{1}{d_3} \sum_{k=1}^{d_3} \| \tilde{T}(::, k) \|_* , \quad \text{where} \quad \tilde{T} = \text{dft}(\mathcal{T}, [], 3) \in \mathbb{R}^{d_1 \times d_2 \times d_3} \]

😊 Orientation sensitivity: computed after DFT along the 3-rd orientation

😊 Order limitation: defined only for 3-way tensors

↓ TNN fails to model

Multi-orientational spectral low-rankness for \( K \)-way \((K \geq 3)\) tensors

↓ This work

😊 Defines 2 Orientation Invariant TNNs for \( K \)-way tensors

😊 Applies them to Robust Tensor Decomposition
Motivation

Low-tubal-rank Structure

Weaknesses of TNN

$$\| T \|_* = \frac{1}{d_3} \sum_{k=1}^{d_3} \| \tilde{T}(::, k) \|_*, \quad \text{where} \quad \tilde{T} = \text{dft}(T, [], 3) \in \mathbb{R}^{d_1 \times d_2 \times d_3}$$

😊 **Orientation sensitivity**: computed after DFT along the 3-rd orientation

😊 **Order limitation**: defined only for 3-way tensors

⇓ TNN fails to model

Multi-orientational spectral low-rankness for $K$-way ($K \geq 3$) tensors

⇓ This work

😊 Defines 2 Orientation Invariant TNNs for $K$-way tensors

😊 Applies them to Robust Tensor Decomposition
### Motivation

**Low-tubal-rank Structure**

#### Weaknesses of TNN

\[
\|\mathcal{T}\|_\star = \frac{1}{d_3} \sum_{k=1}^{d_3} \|\tilde{\mathcal{T}}(:, :, k)\|_\star, \quad \text{where} \quad \tilde{\mathcal{T}} = \text{dft}(\mathcal{T}, [], 3) \in \mathbb{R}^{d_1 \times d_2 \times d_3}
\]

- **Orientation sensitivity**: computed after DFT along the 3-rd orientation
- **Order limitation**: defined only for 3-way tensors

\[\Downarrow\] TNN fails to model 

**Multi-orientational spectral low-rankness for \(K\)-way \((K \geq 3)\) tensors**

\[\Downarrow\] This work

- Defines 2 Orientation Invariant TNNs for \(K\)-way tensors
- Applies them to Robust Tensor Decomposition
Motivation

Low-tubal-rank Structure

Weaknesses of TNN

$$\| \mathcal{T} \|_\star = \frac{1}{d_3} \sum_{k=1}^{d_3} \| \tilde{T}(;, ;, k) \|_\star, \quad \text{where } \tilde{T} = \text{dft}(\mathcal{T}, [], 3) \in \mathbb{R}^{d_1 \times d_2 \times d_3}$$

- **Orientation sensitivity**: computed after DFT along the 3-rd orientation

- **Order limitation**: defined only for 3-way tensors

↓ TNN fails to model

Multi-orientational spectral low-rankness for $K$-way ($K \geq 3$) tensors

↓ This work

- Defines 2 Orientation Invariant TNNs for $K$-way tensors
- Applies them to Robust Tensor Decomposition
Weaknesses of TNN

\[ \| \mathcal{T} \|_* = \frac{1}{d_3} \sum_{k=1}^{d_3} \| \widetilde{T}(:, :, k) \|_* , \text{ where } \widetilde{T} = \text{dft}(\mathcal{T}, [], 3) \in \mathbb{R}^{d_1 \times d_2 \times d_3} \]

- **Orientation sensitivity**: computed after DFT along the 3-rd orientation
- **Order limitation**: defined only for 3-way tensors

\[ \downarrow \text{TNN fails to model} \]

Multi-orientational spectral low-rankness for \( K \)-way \( (K \geq 3) \) tensors

\[ \downarrow \text{This work} \]

- Defines 2 Orientation Invariant TNNs for \( K \)-way tensors
- Applies them to Robust Tensor Decomposition
## Weaknesses of TNN

\[
\| \mathcal{T} \|_* = \frac{1}{d_3} \sum_{k=1}^{d_3} \| \mathcal{T}(::k) \|_*, \quad \text{where} \quad \mathcal{T} = \text{dft} (\mathcal{T}, [\,], 3) \in \mathbb{R}^{d_1 \times d_2 \times d_3}
\]

- **Orientation sensitivity**: computed after DFT along the 3-rd orientation
- **Order limitation**: defined only for 3-way tensors

\[\Downarrow\] **TNN fails to model**

**Multi-orientational spectral low-rankness** for \( K \)-way \((K \geq 3)\) tensors

\[\Downarrow\] **This work**

- Defines 2 Orientation Invariant TNNs for \( K \)-way tensors
- Applies them to Robust Tensor Decomposition
Motivation

Low-tubal-rank Structure

Weaknesses of TNN

\[ \| \mathcal{T} \|_* = \frac{1}{d_3} \sum_{k=1}^{d_3} \| \tilde{T}(; ; k) \|_*, \text{ where } \tilde{T} = \text{dft}(\mathcal{T}, [; ; 3]) \in \mathbb{R}^{d_1 \times d_2 \times d_3} \]

⚠️ **Orientation sensitivity**: computed after DFT along the 3-rd orientation

⚠️ **Order limitation**: defined only for 3-way tensors

↓ TNN fails to model

Multi-orientational spectral low-rankness for \( K \)-way \( (K \geq 3) \) tensors

↓ This work

😊 **Defines 2 Orientation Invariant TNNs for \( K \)-way tensors**

😊 **Applies them to Robust Tensor Decomposition**
### Weaknesses of TNN

\[
\|\mathcal{T}\|_* = \frac{1}{d_3} \sum_{k=1}^{d_3} \|\tilde{\mathcal{T}}(:, :, k)\|_*, \quad \text{where} \quad \tilde{\mathcal{T}} = \text{dft} (\mathcal{T}, [], 3) \in \mathbb{R}^{d_1 \times d_2 \times d_3}
\]

- **Orientation sensitivity**: computed after DFT along the 3-rd orientation
- **Order limitation**: defined only for 3-way tensors

\[\Downarrow \] TNN fails to model

**Multi-orientational spectral low-rankness for K-way \((K \geq 3)\) tensors**

\[\Downarrow \] This work

- Defines 2 Orientation Invariant TNNs for \(K\)-way tensors
- Applies them to Robust Tensor Decomposition
Exploiting multi-orientational spectral low-rankness

Idea: convert a $K$-way tensor to $K$ 3-way tensors then, each 3-way tensor handles one orientation

Step 1: Define mode-$(k, t)$ 3d-unfolding

Step 2: Let $t = k + 1$. Then mode $t$ traverses all the $K$ orientations when $k = 1 : K$.

Step 3: Let $\mathcal{T}_{[k]}$ be the mode-$(k, k + 1)$ 3d-unfolding of $\mathcal{T}$, and use TNN to exploit its spectral low-rankness.
Exploiting multi-orientational spectral low-rankness

**Idea:** convert a $K$-way tensor to $K$ 3-way tensors

then, each 3-way tensor handles one orientation

**Step 1:** Define mode-$(k, t)$ 3d-unfolding

**Step 2:** Let $t = k + 1$. Then mode $t$ traverses all the $K$ orientations when $k = 1 : K$.

**Step 3:** Let $\mathcal{T}_{[k]}$ be the mode-$(k, k+1)$ 3d-unfolding of $\mathcal{T}$, and use TNN to exploit its spectral low-rankness.
Idea: convert a $K$-way tensor to $K$ 3-way tensors
then, each 3-way tensor handles one orientation

Step 1: Define mode-$(k, t)$ 3d-unfolding

Step 2: Let $t = k + 1$. Then mode $t$ traverses all the $K$ orientations when $k = 1 : K$.

Step 3: Let $\mathcal{T}_{[k]}$ be the mode-$(k, k+1)$ 3d-unfolding of $\mathcal{T}$, and use TNN to exploit its spectral low-rankness.
Exploiting multi-orientational spectral low-rankness

**Idea:** convert a $K$-way tensor to $K$ 3-way tensors
then, each 3-way tensor handles one orientation

**Step 1:** Define mode-$(k, t)$ 3d-unfolding

**Step 2:** Let $t = k + 1$. Then mode $t$ traverses all the $K$ orientations when $k = 1 : K$.

**Step 3:** Let $\mathcal{T}_{[k]}$ be the mode-$(k, k + 1)$ 3d-unfolding of $\mathcal{T}$, and use TNN to exploit its spectral low-rankness.
Exploiting multi-orientational spectral low-rankness

Idea: convert a \( K \)-way tensor to \( K \) 3-way tensors
then, each 3-way tensor handles one orientation

Step 1: Define mode-\((k, t)\) 3d-unfolding

Step 2: Let \( t = k + 1 \). Then mode \( t \) traverses all the \( K \) orientations when \( k = 1 : K \).

Step 3: Let \( \mathcal{T}_{[k]} \) be the mode-\((k, k+1)\) 3d-unfolding of \( \mathcal{T} \), and use TNN to exploit its spectral low-rankness.
Exploiting multi-orientational spectral low-rankness

Idea: convert a $K$-way tensor to $K$ 3-way tensors
then, each 3-way tensor handles one orientation

Step 1: Define mode-$(k, t)$ 3d-unfolding

Step 2: Let $t = k + 1$. Then mode $t$ traverses all the $K$ orientations when $k = 1 : K$.

Step 3: Let $\mathcal{T}_{[k]}$ be the mode-$(k, k + 1)$ 3d-unfolding of $\mathcal{T}$, and use TNN to exploit its spectral low-rankness.
Definition 4 (Overlapped OITNN: Sum of TNNs after unfolding).

OITNN-O of \( T \in \mathbb{R}^{d_1 \times \cdots \times d_K} \) is the sum of \( K \) TNNs after 3-d unfoldings

\[
\| T \|_{*o} := \sum_{k=1}^{K} w_k \| T[k] \|_{*},
\]

with weights \( \sum_k w_k = 1 \).

Figure 1: OITNN-O encourages simultaneous low-tubal-rankness in all orientations.
Definition 5 (Latent OITNN: Sum of TNNs after decomposition).

OITNN-L of $\mathcal{T} \in \mathbb{R}^{d_1 \times \cdots \times d_K}$ is the infimum of sum of $K$ TNNs among all decompositions

$$\| \mathcal{T} \|_{*,\ell} := \inf_{\sum_k \mathcal{L}^{(k)} = \mathcal{T}} \sum_{k=1}^{K} v_k \| \mathcal{L}^{(k)} [k] \|_\star,$$

with weights $\sum_k v_k = 1$.

Figure 2: OITNN-L models $\mathcal{T}$ as sum of $K$ low-tubal-rank tensors $\{ \mathcal{L}^{(k)} \}$
Proposed Models for RTD

Model I: RTD based on OITNN-O

\[
(\hat{\mathcal{L}}_o, \hat{S}_o) \in \arg\min_{\mathcal{L}, \mathcal{S}} \frac{1}{2} \| \mathcal{Y} - \mathcal{L} - \mathcal{S} \|_2^2 + \lambda_o \| \mathcal{L} \|_\infty + \mu_o \| \mathcal{S} \|_1
\]

s.t. \( \| \mathcal{L} \|_\infty \leq \alpha \) ← (incoherence condition)

Model II: RTD based on OITNN-L

\[
(\{\hat{\mathcal{L}}^{(k)}\}, \hat{S}_l) \in \arg\min_{\{\mathcal{L}^{(k)}\}, \mathcal{S}} \frac{1}{2} \| \mathcal{Y} - \mathcal{L} - \mathcal{S} \|_2^2 + \lambda_l \sum_k v_k \| \mathcal{L}^{(k)} \|_\infty + \mu_l \| \mathcal{S} \|_1
\]

s.t. \( \| \mathcal{L}^{(l)} \|_\infty \leq \beta \tilde{d}_k, \forall l \neq k; \sum_k \mathcal{L}^{(k)} \|_\infty \leq \alpha \) ← (incoherence condition)
Proposed Models for RTD

Model I: RTD based on OITNN-O

\[
(\hat{L}_o, \hat{S}_o) \in \arg\min_{L, S} \frac{1}{2} \|Y - L - S\|_F^2 + \lambda_o \|L\|_* + \mu_o \|S\|_1 \\
\text{s.t. } \|L\|_\infty \leq \alpha \quad \left(\text{incoherence condition}\right)
\]

Model II: RTD based on OITNN-L

\[
(\{\hat{L}^{(k)}\}, \hat{S}_l) \in \arg\min_{\{L^{(k)}\}, S} \frac{1}{2} \|Y - L - S\|_F^2 + \lambda \sum_k v_k \|L^{(k)}\|_* + \mu \|S\|_1 \\
\text{s.t. } \|L^{(l)}_{[k]}\| \leq \beta \tilde{d}_k, \quad \forall l \neq k; \quad \sum_k \|L^{(k)}\|_\infty \leq \alpha \quad \left(\text{incoherence condition}\right)
\]
Proposed Models for RTD

Model I: RTD based on OITNN-O

\[(\hat{L}_o, \hat{S}_o) \in \arg\min_{L,S} \frac{1}{2} \|Y - L - S\|_F^2 + \lambda_o \|L\|_\diamond + \mu_o \|S\|_1\]

s.t. \[\|L\|_\infty \leq \alpha \quad \leftarrow \text{(incoherence condition)}\]

Model II: RTD based on OITNN-L

\[\{(\hat{L}^{(k)}_l), \hat{S}_l\} \in \arg\min_{\{L^{(k)}_l\},S} \frac{1}{2} \|Y - L - S\|_F^2 + \lambda \sum_k \nu_k \|L^{(k)}_l\|_\diamond + \mu \|S\|_1\]

s.t. \[\|L^{(l)}_k\| \leq \beta \tilde{d}_k, \quad \forall l \neq k; \quad \|\sum_k L^{(k)}_l\|_\infty \leq \alpha \quad \leftarrow \text{(incoherence condition)}\]
Proposed Models for RTD

Model I: RTD based on OITNN-O

\[(\hat{\mathcal{L}}_o, \hat{\mathcal{S}}_o) \in \arg\min_{\mathcal{L}, \mathcal{S}} \frac{1}{2}\|\mathcal{Y} - \mathcal{L} - \mathcal{S}\|_F^2 + \lambda_o \|\mathcal{L}\|_\infty + \mu_o \|\mathcal{S}\|_1 \]

s.t. \(\|\mathcal{L}\|_\infty \leq \alpha\) ← (incoherence condition)

Model II: RTD based on OITNN-L

\[(\{\hat{\mathcal{L}}(k)\}, \hat{\mathcal{S}}_l) \in \arg\min_{\{\mathcal{L}(k)\}, \mathcal{S}} \frac{1}{2}\|\mathcal{Y} - \mathcal{L} - \mathcal{S}\|_F^2 + \lambda \sum_k \nu_k \|\mathcal{L}(k)\|_\infty + \mu \|\mathcal{S}\|_1 \]

s.t. \(\|\mathcal{L}(l)\|_\infty \leq \beta \tilde{d}_k, \forall l \neq k; \sum_k \mathcal{L}(k)\|_\infty \leq \alpha\) ← (incoherence condition)
Proposed Models for RTD

**Model I: RTD based on OITNN-O**

\[(\hat{L}_o, \hat{S}_o) \in \arg\min_{\mathcal{L}, \mathcal{S}} \frac{1}{2} \| \mathcal{Y} - \mathcal{L} - \mathcal{S} \|_F^2 + \lambda_o \| \mathcal{L} \|_{\infty} + \mu_o \| \mathcal{S} \|_1 \]

\[\text{s.t. } \| \mathcal{L} \|_{\infty} \leq \alpha \quad \leftarrow \text{(incoherence condition)}\]

**Model II: RTD based on OITNN-L**

\[(\{\mathcal{L}^{(k)}\}, \hat{S}_\ell) \in \arg\min_{\{\mathcal{L}^{(k)}\}, \mathcal{S}} \frac{1}{2} \| \mathcal{Y} - \mathcal{L} - \mathcal{S} \|_F^2 + \lambda \sum_{k} v_k \| \mathcal{L}^{(k)} \|_{\infty} + \mu \| \mathcal{S} \|_1 \]

\[\text{s.t. } \| \mathcal{L}^{(l)} \|_{\infty} \leq \beta \tilde{d}_k, \quad \forall l \neq k; \quad \sum_k \mathcal{L}^{(k)} \|_{\infty} \leq \alpha \quad \leftarrow \text{(incoherence condition)}\]
Bounds on the Estimation Error

When noise tensor $\mathcal{E}$ has i.i.d. $\mathcal{N}(0, \sigma^2)$ entries

For $\mathcal{L}^* \in \mathbb{R}^{d \times d \times \cdots \times d}$, it holds w.h.p. after parameter tuning:

$$
\frac{\|\hat{\mathcal{L}}_o - \mathcal{L}^*\|_F^2 + \|\hat{\mathcal{S}}_o - \mathcal{S}^*\|_F^2}{d^K} \lesssim \sigma^2 \left( d^{-1} K^{-2} \sum_k r_{tb}(\mathcal{L}^*_k) + \|\mathcal{S}^*\|_{l_0} K \log d \right) \quad \leftarrow \text{(Model I)}
$$

$$
\frac{\|\sum_k \hat{\mathcal{L}}(k) - \mathcal{L}^*\|_F^2 + \|\hat{\mathcal{S}}_e - \mathcal{S}^*\|_F^2}{d^K} \lesssim \sigma^2 \left( d^{-1} \min_k \{ r_{tb}(\mathcal{L}^*_k) \} + \|\mathcal{S}^*\|_{l_0} K \log d \right) \quad \leftarrow \text{(Model II)}
$$

✓ Bound on Model I: controlled by spectral low-rankness of all orientations
✓ Bound on Model II: controlled by the orientation with lowest rank in spectral domain
Bounds on the Estimation Error

When noise tensor $\mathcal{E}$ has i.i.d. $\mathcal{N}(0, \sigma^2)$ entries

For $\mathcal{L}^* \in \mathbb{R}^{d \times d \times \cdots \times d}$, it holds w.h.p. after parameter tuning:

$$\frac{\|\hat{\mathcal{L}}_o - \mathcal{L}^*\|_F^2 + \|\hat{\mathcal{S}}_o - \mathcal{S}^*\|_F^2}{d^K} \leq \sigma^2 \left( d^{-1} K^{-2} \sum_k r_{tb}(\mathcal{L}^*_k) + \|\mathcal{S}^*\|_{l_0} K \log d \right) \quad \leftarrow \text{(Model I)}$$

$$\frac{\|\sum_k \hat{\mathcal{L}}_k - \mathcal{L}^*\|_F^2 + \|\hat{\mathcal{S}}_o - \mathcal{S}^*\|_F^2}{d^K} \leq \sigma^2 \left( d^{-1} \min_k \{ r_{tb}(\mathcal{L}^*_k) \} + \|\mathcal{S}^*\|_{l_0} K \log d \right) \quad \leftarrow \text{(Model II)}$$

- Bound on Model I: controlled by spectral low-rankness of all orientations
- Bound on Model II: controlled by the orientation with lowest rank in spectral domain

---

(NJUST&RIKEN TLU)

OITNN for RTD
Bounds on the Estimation Error

When noise tensor $\mathcal{E}$ has i.i.d. $\mathcal{N}(0, \sigma^2)$ entries

For $\mathcal{L}^* \in \mathbb{R}^{d \times d \times \cdots \times d}$, it holds w.h.p. after parameter tuning:

$$\frac{\|\hat{\mathcal{L}}_o - \mathcal{L}^*\|_F^2 + \|\hat{\mathcal{S}}_o - \mathcal{S}^*\|_F^2}{dK} \lesssim \sigma^2 (d^{-1} K^{-2} \sum_k r_{tb}(\mathcal{L}^*_{[k]}) + \|\mathcal{S}^*\|_{l_0} K \log d) \quad \leftarrow \text{(Model I)}$$

$$\frac{\|\sum_k \hat{\mathcal{L}}(k) - \mathcal{L}^*\|_F^2 + \|\hat{\mathcal{S}} - \mathcal{S}^*\|_F^2}{dK} \lesssim \sigma^2 (d^{-1} \min_k r_{tb}(\mathcal{L}^*_{[k]}) + \|\mathcal{S}^*\|_{l_0} K \log d) \quad \leftarrow \text{(Model II)}$$

✓ Bound on Model I: controlled by spectral low-rankness of all orientations
✓ Bound on Model II: controlled by the orientation with lowest rank in spectral domain
Bounds on the Estimation Error

When noise tensor $\mathcal{E}$ has i.i.d. $\mathcal{N}(0, \sigma^2)$ entries

For $\mathcal{L}^* \in \mathbb{R}^{d \times d \times \cdots \times d}$, it holds w.h.p. after parameter tuning:

For Model I:

$$
\frac{\left\| \hat{L}_o - \mathcal{L}^* \right\|_F^2 + \left\| \hat{S}_o - \mathcal{S}^* \right\|_F^2}{d^K} \lesssim \sigma^2 \left( d^{-1} K^{-2} \sum_k r_{tb}(\mathcal{L}^*_k) \right) + \left\| \mathcal{S}^* \right\|_0 K \log d
$$

For Model II:

$$
\frac{\left\| \sum_k \hat{L}^{(k)} - \mathcal{L}^* \right\|_F^2 + \left\| \hat{S}_l - \mathcal{S}^* \right\|_F^2}{d^K} \lesssim \sigma^2 \left( d^{-1} \min_k \{ r_{tb}(\mathcal{L}^*_k) \} \right) + \left\| \mathcal{S}^* \right\|_0 K \log d
$$

✓ Bound on Model I: controlled by spectral low-rankness of all orientations
✓ Bound on Model II: controlled by the orientation with lowest rank in spectral domain
Bounds on the Estimation Error

When noise tensor $\mathcal{E}$ has i.i.d. $\mathcal{N}(0, \sigma^2)$ entries

For $\mathcal{L}^* \in \mathbb{R}^{d \times d \times \cdots \times d}$, it holds w.h.p. after parameter tuning:

$$\frac{\|\hat{\mathcal{L}}_o - \mathcal{L}^*\|_F^2 + \|\hat{\mathcal{S}}_o - \mathcal{S}^*\|_F^2}{d^K} \lesssim \sigma^2 (d^{-1} K^{-2} \sum_k r_{tb}(\mathcal{L}^*_k) + \|\mathcal{S}^*\|_{l_0} K \log d) \quad \leftarrow \text{(Model I)}$$

$$\frac{\|\sum_k \hat{\mathcal{L}}(k) - \mathcal{L}^*\|_F^2 + \|\hat{\mathcal{S}}_l - \mathcal{S}^*\|_F^2}{d^K} \lesssim \sigma^2 (d^{-1} \min_k \{ r_{tb}(\mathcal{L}^*_k) \} + \|\mathcal{S}^*\|_{l_0} K \log d) \quad \leftarrow \text{(Model II)}$$

✓ Bound on Model I: controlled by spectral low-rankness of all orientations
✓ Bound on Model II: controlled by the orientation with lowest rank in spectral domain

(NJUST&RIKEN TLU)
Bounds on the Estimation Error

When noise tensor $\mathcal{E}$ has i.i.d. $\mathcal{N}(0, \sigma^2)$ entries, for $\mathcal{L}^* \in \mathbb{R}^{d \times d \times \cdots \times d}$, it holds w.h.p. after parameter tuning:

$$\frac{\|\hat{\mathcal{L}}_o - \mathcal{L}^*\|_F^2 + \|\hat{\mathcal{S}}_o - \mathcal{S}^*\|_F^2}{d^K} \lesssim \sigma^2 (d^{-1} K^{-2} \sum_k r_{tb}(\mathcal{L}^*_k) + \|\mathcal{S}^*\|_{l_0} K \log d) \leftarrow \text{(Model I)}$$

$$\frac{\|\sum_k \hat{\mathcal{L}}(k) - \mathcal{L}^*\|_F^2 + \|\hat{\mathcal{S}}_o - \mathcal{S}^*\|_F^2}{d^K} \lesssim \sigma^2 (d^{-1} \min_k \{r_{tb}(\mathcal{L}^*_k)\} + \|\mathcal{S}^*\|_{l_0} K \log d) \leftarrow \text{(Model II)}$$

✓ Bound on Model I: controlled by spectral low-rankness of all orientations
✓ Bound on Model II: controlled by the orientation with lowest rank in spectral domain
Robust Image Recovery

Figure 3: Robust image recovery with different corruption ratio $s$ and noise level $c$.

(a) $(s, c) = (0.05, 0.1)$

(b) $(s, c) = (0.15, 0.15)$
Experiments

Image Completion

1. Setting I: 90% random missing
2. Setting II: rows and columns missing, total ratio 85%

Figure 4: Quantitative comparison in image completion.
Experiments

1. Row 1: robust image recovery with corruption ratio $s = 0.05$ and noise level $c = 0.1$
2. Row 2: image completion with 90% random missing entries
3. Row 3: image completion with missing columns and rows (total missing ratio 85%)
Row 1: robust image recovery with corruption ratio $\delta = 0.05$ and noise level $c = 0.1$

Row 2: image completion with 90% random missing entries

Row 3: image completion with missing columns and rows (total missing ratio 85%)
1. Robust image recovery with corruption ratio $s = 0.05$ and noise level $c = 0.1$.
2. Image completion with 90% random missing entries.
3. Image completion with missing columns and rows (total missing ratio 85%).
Row 1: robust image recovery with corruption ratio $s = 0.05$ and noise level $c = 0.1$
Row 2: image completion with 90% random missing entries
Row 3: image completion with missing columns and rows (total missing ratio 85%)
Experiments

Video Completion

Figure 5: Video completion with 90% random missing
Conclusion

Contributions

1. We defined two new norms for $K$-way ($K \geq 3$) tensors.
2. We presented two models for RTD with error bounds.

Thank you.
Conclusion

Contributions

1. We defined two new norms for $K$-way ($K \geq 3$) tensors.
2. We presented two models for RTD with error bounds.

Thank you.
Conclusion

Contributions

1. We defined two new norms for $K$-way ($K \geq 3$) tensors.
2. We presented two models for RTD with error bounds.

Thank you.