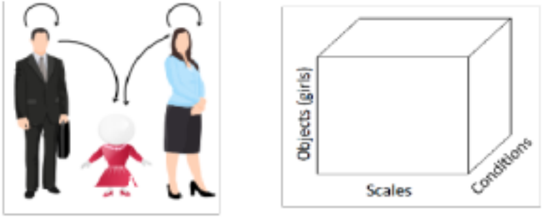
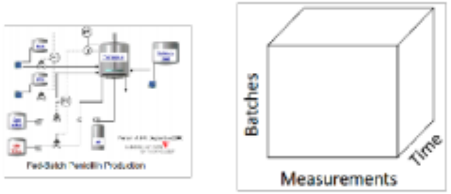
An L1-L2 Variant of Tubal Nuclear Norm for Guaranteed Tensor Recovery

Andong Wang, Guoxu Zhou, Zhong Jin, Qibin Zhao

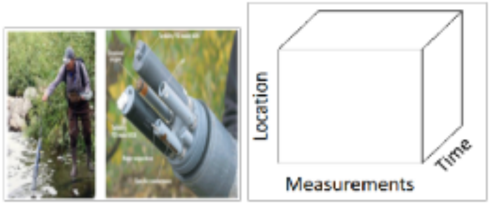
Tensor data is almost everywhere!



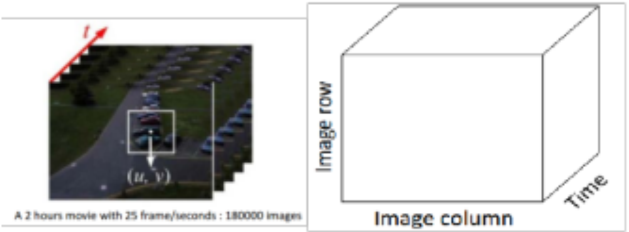
Psychology
Behavior analysis



Process monitoring
Failure detection

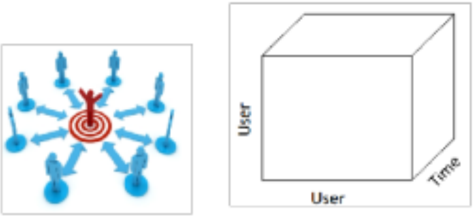


Environment monitoring
Quality assessment

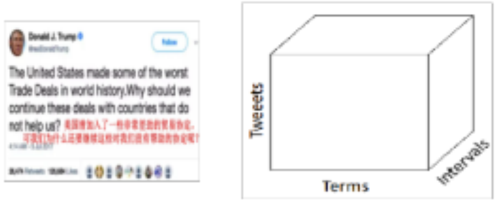


Video surveillance
Anomaly detection

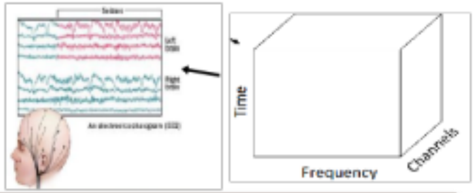
Image/Video processing
Inpainting/De-noising



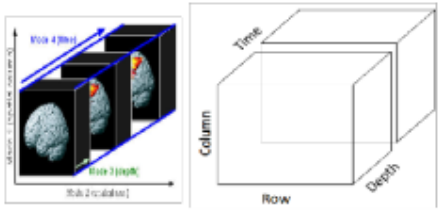
Social networks
Community detection



Question system
Topic model



EEG signal processing
Disease surveillance



MRI
Behavior recognition

Tensor Recovery

Tensor data is almost everywhere!

But in many applications, only **a few noisy observations** are available.

The Observation Model

$$\overbrace{y_i}^{\text{observation}} = \langle \underbrace{\mathcal{L}^*}_{\text{true tensor}}, \overbrace{\mathcal{X}_i}^{\text{design tensor}} \rangle + \underbrace{\xi_i}_{\text{noise}}, \quad i = 1, \cdots, N$$

How to **recover the true tensor** from a few noisy observations?

To recover $\mathcal{L}^*$, the **tubal nuclear norm** based models were proposed.

C. Lu et al. Exact low tubal rank tensor recovery from Gaussian measurements. IJCAI 2018.

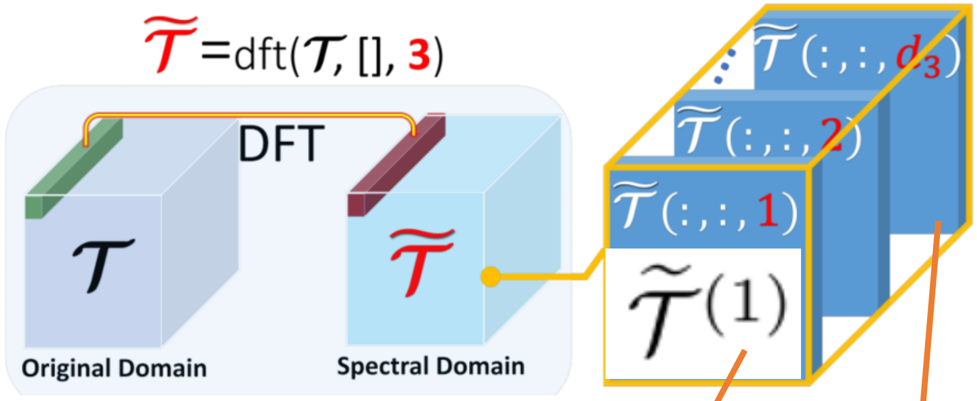
Tubal nuclear norm (TNN)

TNN is the averaged nuclear norm of frontal slices after DFT

$$\|\mathcal{T}\|_{\circledast} := \frac{1}{d_3}\sum_{i=1}^{d_3}\|\widetilde{\mathcal{T}}^{(i)}\|_*$$

also equal to the scaled nuclear norm of **Fourier block diagonal matrix**

$$\|\mathcal{T}\|_{\circledast} = \frac{1}{d_3}\|\overline{\mathbf{T}}\|_*$$



$$\overline{\mathbf{T}} := \mathrm{blkdiagDFT}(\mathcal{T}) = \begin{bmatrix} \widetilde{\mathcal{T}}^{(1)} & & \\ & \ddots & \\ & & \widetilde{\mathcal{T}}^{(d_3)} \end{bmatrix}$$

TNN yields sub-optimal performance due to ***biased approximation of*** $\mathrm{rank}(\overline{\mathbf{T}})$

T. Ji et al. A new surrogate for tensor multirank and applications in image and video completion, IEEE-PIC 2017.

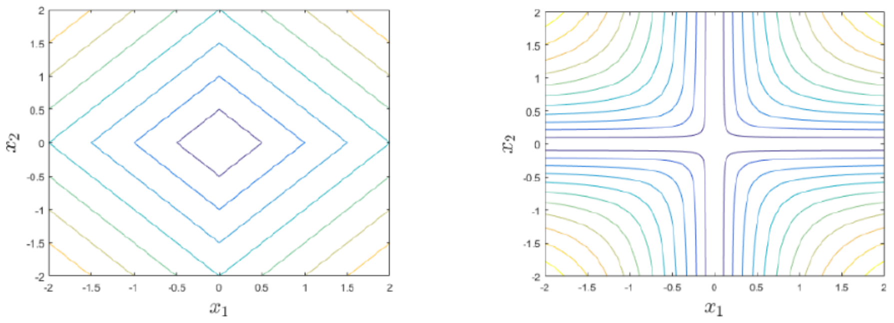
Tensor L1-L2 metric for tensor recovery

L1-L2 metric

With $\alpha > 0$, define the vector $l_1 - \alpha l_2$ metric

$$\|\mathbf{x}\|_{\alpha,1-2} := \|\mathbf{x}\|_1 - \alpha\|\mathbf{x}\|_2.$$

A tighter approximation of L0-norm than L1-norm



(a) l_1-norm (b) $l_1 - \alpha l_2$ metric ($\alpha = 1$)

Figure 1: Level curves of the l_1-norm and $l_1 - \alpha l_2$ metric ($\alpha = 1$).

[1] Y. Lou et al. Fast l1–l2 minimization via a proximal operator. Journal of Scientific Computing 2018.

[2] Q. Yao et al. Fast learning with nonconvex l1-2 regularization. arXiv 2016.

Tensor L1-L2 metric

singular values

$$\|\mathcal{T}\|_{\alpha,\circledast-\mathrm{F}} := \frac{1}{d_3}\|\boxed{\boldsymbol{\sigma}(\overline{\mathbf{T}})}\|_{\alpha,1-2}$$

A tighter approximation of $\mathrm{rank}(\overline{\mathbf{T}})$ than TNN

The proposed estimator

$$\hat{\mathcal{L}} \in \operatorname*{argmin}_{\mathcal{L}} \|\mathcal{L}\|_{\alpha,\circledast-\mathrm{F}}$$

$$\text{s.t. } \|\mathbf{y} - \mathfrak{X}(\mathcal{L})\| \le \tau$$

Bound on estimation error

"**How well** can we estimate the true tensor?"

bound the estimation error

$$\|\hat{\mathcal{L}} - \mathcal{L}^*\|_{\mathrm{F}}$$

Theorem 2 (Stable recovery under tm-RIP). *Suppose the true tensor $\mathcal{L}^*$ in Model (4) has multi-rank $\mathbf{r} = (r_1, \cdots, r_{d_3})$. If there exists a positive integer s such that*

$$\Phi_{\mathbf{r},s} := 1 - \delta^{\mathfrak{X}}_{2\mathbf{r}+\mathbf{s}} - \frac{\sqrt{4\|\mathbf{r}\|_1 + 2\alpha^2 d_3}}{\sqrt{s} - \alpha}(\delta^{\mathfrak{X}}_{2\mathbf{r}+\mathbf{s}} + \delta^{\mathfrak{X}}_{2\mathbf{s}}) > 0, \quad (8)$$

where $\mathbf{s} = (s, \cdots, s) \in \mathbb{R}^{d_3}$, then any $\hat{\mathcal{L}}$ in Eq. (5) obeys

$$\|\hat{\mathcal{L}} - \mathcal{L}^*\|_{\mathrm{F}} \leq \frac{2\sqrt{1 + \delta^{\mathfrak{X}}_{2\mathbf{r}+\mathbf{s}}}}{\Phi_{\mathbf{r},s}}\tau. \quad (9)$$

Under a new **tensor RIP** condition, we guarantee **stable tensor recovery.**

The optimization algorithm

"How to **solve the proposed model?**"

1

Derive a **closed-from solution** of the **proximality operator** the metric.

2

Add auxiliary variables for better decoupling, and use **ADMM**

$$
\hat{\mathcal{L}} \in \operatorname*{argmin}_{\mathcal{L}} \|\mathcal{L}\|_{\alpha,\circledast-\mathrm{F}} \quad \text{s.t.} \quad \|\mathbf{y} - \mathfrak{X}(\mathcal{L})\| \le \tau
$$

auxiliary variables

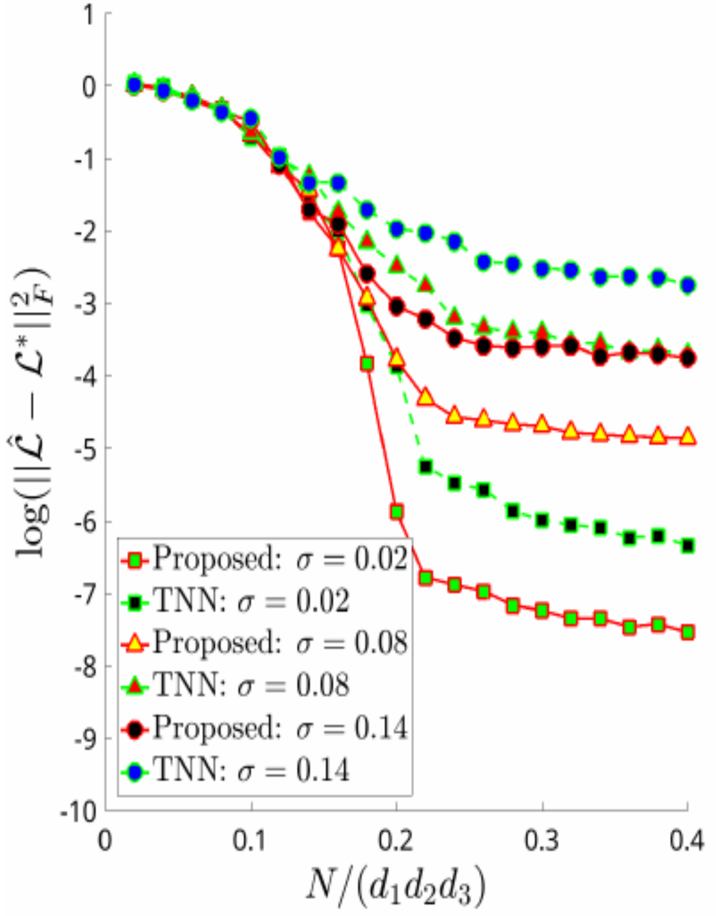
$$
\min_{\mathcal{L},\mathcal{K},\boldsymbol{\epsilon}} \|\mathcal{L}\|_{\alpha,\circledast-\mathrm{F}} \quad \text{s.t.} \quad \mathcal{K} = \mathcal{L},\ \mathfrak{X}(\mathcal{K}) + \boldsymbol{\epsilon} = \mathbf{y}, \boldsymbol{\epsilon} \in \mathbb{B}_{\tau}.
$$

[1] Y. Lou et al. Fast l1–l2 minimization via a proximal operator. Journal of Scientific Computing 2018.
[2] Q. Yao et al. Fast learning with nonconvex l1-2 regularization. arXiv 2016.

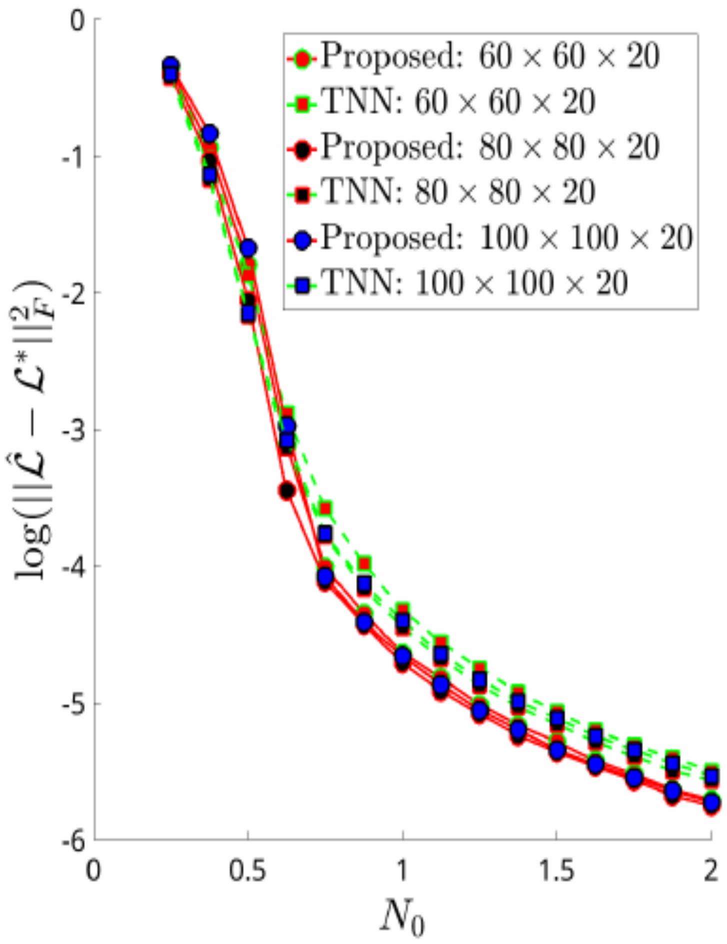
Experimental results

Synthetic data

tensor compressive sensing

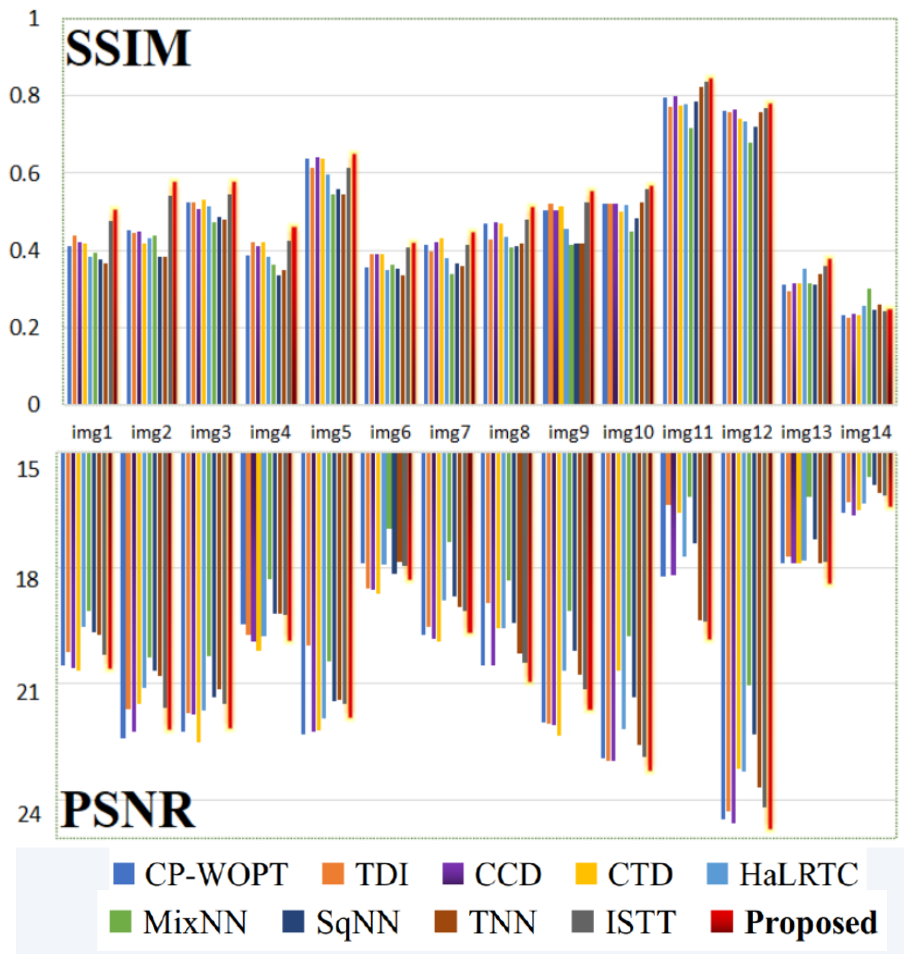


tensor completion



Real data

image inpainting



The proposed metric has promising performance in tensor recovery.


THANKS