ROBUST LOW-TUBAL-RANK TENSOR COMPLETION

Andong Wang[1,2,3], *Xulin Song*[1,2], *Xiyin Wu*[1,2], *Zhihui Lai*[4] *and Zhong Jin*[1,2*]

[1]School of Computer Science and Engineering, Nanjing University of Science & Technology, China
[2]Key Laboratory of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education, Nanjing University of Science & Technology, China
[3] Tensor Learning Unit, RIKEN Center for Advanced Intelligence Project (AIP), Japan
[4]School of Computer Science and Software Engineering, Shenzhen University, China

ABSTRACT

Real multi-way data may suffer from missing entries, noise and outliers simultaneously. The recently proposed tubal nuclear norm (TNN) has shown its superiority in tensor completion. However, statistical analysis of TNN based models is still deficient. This paper aims to robustly recover a polluted incomplete tensor with rigorous statistical guarantee. Specifically, an estimator based on a weighed variant of TNN is proposed to complete a low-tubal-rank tensor corrupted by element sparse errors or slice sparse sample outliers from partial noisy observations. Non-asymptotic upper bounds on the estimation error are established and further proved to be minimax optimal up to a log factor. Sharpness of the upper bounds is verified on synthetic datasets and superiority of the proposed estimator is demonstrated through robust video inpainting.

Index Terms— robust tensor completion, tensor SVD, tubal rank, statistical performance

1. INTRODUCTION

To recover a multi-way array from partial observations, tensor completion has been an active topic in signal processing, machine learning, computer vision, etc [1–4]. Due to various reasons such as partial sensor failures, communication errors, and occlusion by obstacles, the data tensor may face missing entries, noise and outliers at the same time [5–7]. In these circumstances, one needs to complete the tensor robustly.

Obviously, it is impossible to complete a tensor from partial and corrupted observations unless some additional assumptions are made. In this paper, we assume that the underlying tensor is low-tubal-rank. The tensor tubal rank is defined through the tensor singular value decomposition (t-SVD) which is (to the best of our knowledge) the only known multi-linear extension of matrix SVD that has best rank-k approximation guarantee (like the Eckart-Young theorem) [8]. As pointed out by [9, 10], the low-tubal-rank model is ideal for capturing the "spatial-shifting" property which is ubiquitous in real data arrays. As the most representative low-tubal-rank model, the tubal nuclear norm (TNN) minimization based method has shown superiority over traditional tensor low-rank models in many tensor recovery problems like tensor completion [11–13], tensor robust principle component analysis (TRPCA) [14], outlier robust tensor PCA (OR-TPCA) [15].

In this paper, we study robust low-tubal-rank tensor completion which aims to recover a sparsely corrupted low-tubal-rank tensor from its partial noisy observations. To estimate the underlying tensor, an estimator based on a slice weighed variant of TNN named SwTNN is defined. Then, we establish non-asymptotic upper bounds on the estimation error and prove that the upper bounds are minimax optimal up to a log factor. We verify the sharpness of the proposed error bounds through numerical experiments. Effectiveness of the estimator is evaluated by experiments on real-world datasets.

2. PRELIMINARIES AND THE PROPOSED NORM

2.1. Notations

Main notations are listed in Table 1. Let $[d] := \{1, \cdots, d\}$, $\forall d \in \mathbb{N}_+$. Let $a \vee b = \max\{a, b\}$ and $a \wedge b = \min\{a, b\}, \forall a, b \in \mathbb{R}$. For $i \in [d], \boldsymbol{e}_i \in \mathbb{R}^d$ denotes the standard vector basis whose i_{th} entry is 1 with the others 0. For $(i, j, k) \in [d_1] \times [d_2] \times [d_3]$, the outer product $\boldsymbol{e}_i \circ \boldsymbol{e}_j \circ \boldsymbol{e}_k$ denotes a standard tensor basis in $\mathbb{R}^{d_1 \times d_2 \times d_3}$. For a 3-way tensor, fft3$(\cdot)$ and ifft3$(\cdot)$ denotes the *fast Fourier transformation* and the *inverse fast Fourier transformation* along *the third mode*. For a set Ψ, $|\Psi|$ denotes its cardinality and $\Psi^\perp$ its complement. Positive constants are denoted by C, c, c_0, etc. When the field and size of a tensor are not shown explicitly, it is in $\mathbb{R}^{d_1 \times d_2 \times d_3}$. The spectral norm $\|\cdot\|$ and nuclear norm $\|\cdot\|_*$ of a matrix are the maximum and the sum of the singular values, respectively.

2.2. T-SVD and SwTNN

First, we define some concepts of t-SVD.

*✉ *zhongjin@njust.edu.cn*. This work is partially supported by National Natural Science Foundation of China under Grant Nos. 61872188, U1713208, 61602244, 61672287, 61702262, 61773215, 61703209, and by the Natural Science Foundation of Guangdong Province under Grant No. 2017A030313367.

Table 1: List of notations

Notations	Descriptions	Notations	Descriptions
$\mathcal{L}^*$	true low-rank tensor	$\mathcal{S}^*$	true "sparse" tensor
$\widetilde{\mathcal{L}}$	fft3($\mathcal{L}$)	$\|\mathcal{L}\|_0$	$\sum_{ijk} \mathbf{1}(\mathcal{L}_{ijk} \neq 0)$
$\mathcal{L}_{ijk}$	$(i,j,k)_{th}$ entry of $\mathcal{L}$	$\|\mathcal{L}\|_1$	$\sum_{ijk} \|\mathcal{L}_{ijk}\|$
$\mathcal{L}(i,j,:)$	$(i,j)_{th}$ tube of $\mathcal{L}$	$\|\mathcal{L}\|_{\mathrm{F}}$	$\sqrt{\sum_{ijk} \mathcal{L}_{ijk}^2}$
$\mathcal{L}(:,j,:)$	j_{th} lateral slice of $\mathcal{L}$	$\|\mathcal{L}\|_\infty$	$\max_{ijk} \|\mathcal{L}_{ijk}\|$
$\mathcal{L}(:,:,k)$	k_{th} frontal slice of $\mathcal{L}$	$\|\mathcal{L}\|_{\text{slice},1}$	$\sum_j \|\mathcal{L}(:,j,:)\|_{\mathrm{F}}$
Θ_s	support of $\mathcal{S}^*$	$\|\mathcal{L}\|_{\text{slice},0}$	$\sum_j \mathbf{1}(\mathcal{L}(:,j,:) \neq \mathbf{0})$
$\Theta_s^\perp$	complement of Θ_s	$\|\mathcal{L}\|_{\text{slice},\infty}$	$\max_j \|\mathcal{L}(:,j,:)\|_{\mathrm{F}}$
$\boldsymbol{\mathcal{A}}(r,\gamma_0,\alpha) := \{(\mathcal{L},\mathcal{S}) : r_t(\mathcal{L}) \le r, \gamma(\mathcal{S}) \le \gamma_0, (\|\mathcal{L}\|_\infty \vee \|\mathcal{S}\|_\infty) \le \alpha\}$			

Definition 1 (t-product [16]) *Let $\mathcal{T}_1 \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ and $\mathcal{T}_2 \in \mathbb{R}^{d_2 \times d_4 \times d_3}$. Their t-product*

$$\mathcal{T} := \mathcal{T}_1 * \mathcal{T}_2$$

is a tensor in $\mathbb{R}^{d_1 \times d_4 \times d_3}$, whose $(i,j)_{th}$ tube $\mathcal{T}(i,j,:) = \sum_{k=1}^{d_2} \mathcal{T}_1(i,k,:) \bullet \mathcal{T}_2(k,j,:)$, where $\bullet$ denotes the circular convolution [8].

Based on the tensor transpose, f-diagonal tensor and orthogonal tensor [16], the t-SVD can be defined.

Definition 2 (t-SVD, Tensor tubal rank [16]) *Any tensor $\mathcal{T}$ has tensor singular value decomposition (t-SVD) as follows*

$$\mathcal{T} := \mathcal{U} * \underline{\mathbf{\Lambda}} * \mathcal{V}^\top,$$

where $\mathcal{U} \in \mathbb{R}^{d_1 \times d_1 \times d_3}$ and $\mathcal{V} \in \mathbb{R}^{d_2 \times d_2 \times d_3}$ are orthogonal, $\underline{\mathbf{\Lambda}} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ is f-diagonal, $(\cdot)^\top$ denotes the tensor transpose. The tensor tubal rank of $\mathcal{T}$ is defined as the number of non-zero tubes of $\underline{\mathbf{\Lambda}}$ in its t-SVD, i.e., $r_t(\mathcal{T}) := \sum_i \mathbf{1}(\underline{\mathbf{\Lambda}}(i,i,:) \neq \mathbf{0})$.

Motivated by the advantage of matrix weighted nuclear norm [17] over the nuclear norm, we define the the slice-weighted TNN (SwTNN).

Definition 3 (Slice-weighted TNN) *The slice weighted tubal nuclear norm (SwTNN) of $\mathcal{T} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ is defined as*

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

where $\widetilde{\mathcal{T}} :=$ fft3($\mathcal{T}$), and w_k's are positive parameters. If all w_k's are equal 1, SwTNN reduces to the tubal nuclear norm (TNN) [14].

The superiority of SwTNN over TNN lies in that w_k's in Eq. (1) can provide flexibility in emphasizing the effect of different frequency components $\widetilde{\mathcal{T}}(:,:,k)$ of the signal tensor. The following properties of SwTNN are introduced:

Lemma 1 *SwTNN is a valid norm with dual norm*

$$\|\mathcal{T}\|_{w\star}^* := \sup_{\|\mathcal{X}\|_{w\star} \le 1} \langle \mathcal{T}, \mathcal{X} \rangle = \max_k \left\{ \|\widetilde{\mathcal{T}}(:,:,k)\| / w_k \right\}.$$

Lemma 2 *Let $\mathcal{T}_0$ have t-SVD: $\mathcal{T}_0 = \mathcal{U}_0 * \underline{\mathbf{\Lambda}}_0 * \mathcal{V}_0{}^\top$. The proximal operator of SwTNN at $\mathcal{T}_0$ defined by $\mathrm{Prox}_\tau^{\|\cdot\|_{w\star}}(\mathcal{T}_0) := \operatorname{argmin}_{\mathcal{T}} \tau \|\mathcal{T}\|_{w\star} + \frac{1}{2}\|\mathcal{T} - \mathcal{T}_0\|_F^2$ can be computed as*

$$\mathrm{Prox}_\tau^{\|\cdot\|_{w\star}}(\mathcal{T}_0) = \mathcal{U}_0 * \mathrm{ifft3}(\widetilde{\underline{\mathbf{\Lambda}}}) * \mathcal{V}_0{}^\top,$$

where $\widetilde{\underline{\mathbf{\Lambda}}}(:,:,k) = \max\{\widetilde{\underline{\mathbf{\Lambda}}}_0(:,:,k) - w_k\tau, 0\}$, for all $k \in [d_3]$.

3. ROBUST TENSOR COMPLETION

3.1. The Observation Model

Let $\mathcal{L}^*$ denote the true "signal" tensor with low tubal rank, i.e., $r_t(\mathcal{L}^*) \ll d_1 \wedge d_2$. Let $\mathcal{S}^*$ denote the "corruption" tensor with support Θ_s. We suppose $\mathcal{S}^*$ belongs to one of the two settings: the TRPCA [14] setting where $\mathcal{S}^*$ carries *element sparse errors* (i.e. $\|\mathcal{S}^*\|_0 \ll d_1 d_2 d_3$), or the OR-TPCA[15] setting where $\mathcal{S}^*$ models *slice sparse sample outliers* (i.e., $\|\mathcal{S}^*\|_{\text{slice},0} \ll d_2$).

Suppose we observe N scalars y_i from the model

$$y_i = \langle \mathcal{L}^* + \mathcal{S}^*, \mathcal{X}_i \rangle + \xi_i, \ \ \forall i \in [N], \tag{2}$$

where $\xi_i \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$ with known σ and $\mathcal{X}_i \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ are the design tensors satisfying the following assumptions:

Assumption 1 *The following assumptions are made:*

(I) First, all corrupted positions are observed, i.e., the support of $\mathcal{S}^*$ is totally observed[1]. Or formally, there exists an unknown subset[2] $\boldsymbol{\mathcal{X}}_s \subset \{\mathcal{X}_i\}_{i=1}^N$ sampled from any distribution $\mathbf{\Pi}_{\Theta_s}$ on the set $\boldsymbol{\mathcal{X}}_{\Theta_s} := \{\boldsymbol{e}_i \circ \boldsymbol{e}_j \circ \boldsymbol{e}_k, \forall (i,j,k) \in \Theta_s\}$, such that each element in $\boldsymbol{\mathcal{X}}_{\Theta_s}$ is sampled at least once.

(II) Second, all non-corrupted positions are sampled uniformly with replacement. Each element of the set $\boldsymbol{\mathcal{X}}_s^\perp := \{\mathcal{X}_i\}_{i=1}^N \backslash \boldsymbol{\mathcal{X}}_s$ is sampled i.i.d. from uniform distribution $\mathbf{\Pi}_{\Theta_s^\perp}$ on the set $\boldsymbol{\mathcal{X}}_{\Theta_s^\perp} := \{\boldsymbol{e}_i \circ \boldsymbol{e}_j \circ \boldsymbol{e}_k, \forall (i,j,k) \in \Theta_s^\perp\}$.

According to the observation model (2) and Assumption 1, the true tensor $\mathcal{L}^*$ is firstly corrupted by sparse $\mathcal{S}^*$ and then sampled with additive Gaussian noise ξ_i's. The corrupted positions of $\mathcal{L}^*$ are assumed to be fully observed with design tensors $\boldsymbol{\mathcal{X}}_s \subset \{\mathcal{X}_i\}_{i=1}^N$, and the remaining non-corrupted positions are sampled uniformly through design tensors $\boldsymbol{\mathcal{X}}_s^\perp := \{\mathcal{X}_i\}_{i=1}^N \backslash \boldsymbol{\mathcal{X}}_s$.

3.2. Problem Formulation

Given partial noisy observations $\{(\mathcal{X}_i, y_i)\}_{i=1}^N$ from observation model (2), the goal of robust tensor completion is to recover $\mathcal{L}^*$ and $\mathcal{S}^*$. The problem can be regarded as a robust variant of tensor completion in [16], a noisy partial variant of TRPCA [14] and OR-TPCA [15], and a noisy variant of [5].

Exploiting the low-tubal-rankness of $\mathcal{L}^*$ and the sparsity of $\mathcal{S}^*$, we consider the following problem:

$$\min_{\mathcal{L},\mathcal{S}} \ l(\mathcal{L},\mathcal{S}) + \lambda_\iota r_t(\mathcal{L}) + \lambda_s R_0(\mathcal{S}), \tag{3}$$

where $\lambda_\iota, \lambda_s \ge 0$ are regularization parameters, $l(\mathcal{L},\mathcal{S}) = \frac{1}{2}\sum_{i=1}^N \left(y_i - \langle \mathcal{L} + \mathcal{S}, \mathcal{X}_i \rangle \right)^2$ is the data fidelity term and $R_0(\mathcal{S})$

[1]This assumption is necessary, because it is impossible to recover the unobserved entries of $\mathcal{S}^*$ [5].

[2]With a slight abuse of notation, the concept 'set' in this section allows multiplicity of elements.

is $\|\cdot\|_0$ or $\|\cdot\|_{\text{slice},0}$ when $\mathcal{S}^*$ represents element sparse errors or slice sparse sample outliers.

Due to the non-convexity of tubal rank $r_t(\cdot)$ and sparsity $R_0(\cdot)$, Problem (3) is NP hard. We use convex relaxations of $r_t(\cdot)$ and $R_0(\cdot)$, i.e., $\|\cdot\|_{w\star}$ and $R(\cdot) \in \{\|\cdot\|_1, \|\cdot\|_{\text{slice},1}\}$, to replace them, respectively. Specifically, we define the following estimator

$$(\hat{\mathcal{L}}, \hat{\mathcal{S}}) := \underset{\substack{\|\mathcal{L}\|_\infty \le \alpha \\ \|\mathcal{S}\|_\infty \le \alpha}}{\operatorname{argmin}} \quad l(\mathcal{L}, \mathcal{S}) + \lambda_\iota \|\mathcal{L}\|_{w\star} + \lambda_s R(\mathcal{S}), \tag{4}$$

where $\alpha > 0$ is an upper estimate of $\|\mathcal{L}^*\|_\infty$ and $\|\mathcal{S}^*\|_\infty$. The additional constraints $\|\mathcal{L}\|_\infty \le \alpha$ and $\|\mathcal{S}\|_\infty \le \alpha$ are introduced to exclude the "spiky" tensors, which is important in controlling the separability of $\mathcal{L}^*$ and $\mathcal{S}^*$. Such "non-spiky" constraints are also imposed in previous literatures [12, 18, 19], playing a key role in bounding the estimation error[3].

4. STATISTICAL PERFORMANCE

Let $N_S = |\boldsymbol{\mathcal{X}}_s|$ and $N_L = |\boldsymbol{\mathcal{X}}_s^\perp|$ denote the number of corrupted and uncorrupted observations, respectively. Let $\gamma(\mathcal{S}^*)$ denote the corruption ratio which equals $\|\mathcal{S}^*\|_0/(d_1 d_2 d_3)$ or $\|\mathcal{S}^*\|_{\text{slice},0}/d_2$ when $\mathcal{S}^*$ represents element sparse errors or slice sparse sample outliers, respectively. Without loss of generality, assume $d_1 \ge d_2$. For simplicity, let $\tilde{d} = (d_1 + d_2)d_3$, $D = d_1 d_2 d_3$, $w_{\max} = \max_k w_k$, $w_{\min} = \min_k w_k$, $\Delta^2_{\iota,s} = \|\hat{\mathcal{L}} - \mathcal{L}^*\|_F^2 + \|\hat{\mathcal{S}} - \mathcal{S}^*\|_F^2$. We establish the following upper bounds on the estimation error.

Theorem 1 (Upper bounds on the estimation error) *Let $N_L \ge c_1 d_1 d_3 \log \tilde{d} \log^2 d_2$. Choose parameters $\lambda_\iota = c_2(\sigma \vee \alpha) w_{\min}^{-1} \sqrt{N \log(\tilde{d})/d_2}$, and $\lambda_s = c_3(\sigma \vee \alpha) log\tilde{d}$ or $c_4(\sigma \vee \alpha)\sqrt{N \log \tilde{d}/d_2}$ for $R(\cdot)$ being $\|\cdot\|_1$ or $\|\cdot\|_{slice,1}$ in Problem (4), respectively. Then it holds with probability at least $1 - c_5/\tilde{d}$:*

$$\Delta^2_{\iota,s}/D \le C(\boldsymbol{E}_L + \boldsymbol{E}_S), \tag{5}$$

where $\boldsymbol{E}_L = (\sigma \vee \alpha)^2 \big(w_{\max} w_{\min}^{-1} r_t(\mathcal{L}^*) d_1 d_3 \log \tilde{d}\big)/N_L$ and $\boldsymbol{E}_S = N_S \log \tilde{d}/N_L + \alpha^2 \gamma(\mathcal{S}^*)$.

The proof follows [12] and [18], and is omitted due to space limitation. According to Theorem 1, if σ, α and w_k's are fixed and each of the corrupted positions are observed exactly once, then the upper bound would (with high probability) scale like

$$O\Big(r_t(\mathcal{L}^*) \cdot \frac{d_1 d_3 \log \tilde{d}}{N_L} + \gamma(\mathcal{S}^*) \cdot \big(\frac{D \log \tilde{d}}{N_L} + 1\big)\Big), \tag{6}$$

where the first error term accounts for tensor completion and the second term stems from corruption.

[3]The constraint is a compromise that $\mathcal{L}^*$ is not assumed to satisfy the stringent tensor coherent condition (TIC), which essentially ensures the separability of $\mathcal{L}^*$ and $\mathcal{S}^*$ [5, 14].

Remark 1 (Connection with previous works) *The proposed bounds consist with previous works:*

(I) If $\gamma(\mathcal{S}^*) = 0$, i.e., the corruption $\mathcal{S}^*$ vanishes, we have $\|\hat{\mathcal{L}} - \mathcal{L}^*\|_F^2/D = O(r_t(\mathcal{L}^*) d_1 d_3 \log \tilde{d}/N)$, which is consistent with the error bound for noisy low-tubal-rank tensor completion in [20, 21].

(II) According to Eq. (6), $r_t(\mathcal{L}^*)$ can take the order $O(d_2/\log \tilde{d})$ and $\gamma(\mathcal{S}^*)$ can be $O(1)$ for approximate estimation with small error. It is consistent (up to a logarithm factor) with the results for exact completion in [5] which ensures $r_t(\mathcal{L}^*) = O(d_2/\log^2 \tilde{d})$ and $\gamma(\mathcal{S}^*) = O(1)$.

(III) When $d_3 = 1$, the error bounds reduce to results for the robust matrix completion [18].

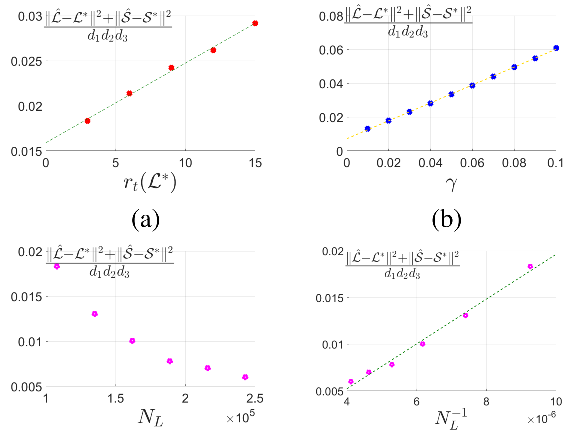
Remark 2 (No exact recovery guarantee) *According to Theorem 1, when $\sigma = 0$ and $\gamma(\mathcal{S}^*) = 0$, i.e., in the noiseless case, the estimation error is upper bounded by $\alpha r_t(\mathcal{L}^*) d_1 d_3 \log \tilde{d}/N$ which is not zero. Thus, no exact recovery is guaranteed. It can be seen as a trade-off that we do not assume the low-tubal-rank tensor $\mathcal{L}^*$ to satisfy the tensor incoherent conditions [5, 14, 15] which essentially ensures the separability between $\mathcal{L}^*$ and any sparse tensors.*

To explore the optimality of the proposed upper bounds, we establish the minimax lower bounds on the estimation error when $(\mathcal{L}^*, \mathcal{S}^*)$ belongs to the tensor class $\boldsymbol{\mathcal{A}}(r, \gamma_0, \alpha)$ (see definition in Table 1). Define $\phi(N, r, \gamma_0) := (\sigma \wedge \alpha)^2 \big(N_L^{-1}(r d_1 d_3 + N - N_L) + \gamma_0\big)$. We have the following theorem.

Theorem 2 (Minimax lower bounds) *Assume that $d_1, d_2 \ge 2$, $\gamma_0 \le 1/2$, $r d_1 d_3 \le N_L$ and there exists a constant $\tau > 0$ such that $N_S \le \tau r \tilde{d}$. Then, there exist absolute constants $\beta \in (0, 1)$ and $c > 0$, such that*

$$\inf_{(\hat{\mathcal{L}}, \hat{\mathcal{S}})} \sup_{\substack{(\mathcal{L}^*, \mathcal{S}^*) \\ \in \boldsymbol{\mathcal{A}}(r, \gamma_0, \alpha)}} \mathbb{P}_{(\mathcal{L}^*, \mathcal{S}^*)}\Big(\frac{\Delta^2_{\iota,s}}{D} \ge c\phi(N, r, \gamma_0)\Big) \ge \beta. \tag{7}$$

The lower bounds indicate that the upper bounds in Theorem 1 are minimax optimal (up to a logarithm factor). In other words, no estimator can achieve better estimation performance than the proposed estimator (up to a logarithm factor) in the case where $\mathcal{L}^*$ and $\mathcal{S}^*$ are in $\boldsymbol{\mathcal{A}}(r, \gamma_0, \alpha)$. The sharpness of the upper bound is verified through simulations in Section 5.1. Note that in the no-corruption case, the upper bound will reduce to $r d_1 d_3/N_L$, indicating that the order of sample complexity N_L should not be lower than $O(r d_1 d_3)$ to obtain an estimation with small error. The order can not be further tightened for general low-tubal-rank tensors, since the degree of freedom of a general 3-way tensor $\mathcal{T} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ with tubal rank r is at most $r d_3(d_1 + d_2 - r)$ [20, 22].



(a) (b) (c) (d)

Fig. 1: Error vs $r = r_t(\mathcal{L}^*)$, γ, N_L and N_L^{-1} : (a) Error vs r with $\gamma = 0.01$ and $N_L = 0.4(1-\gamma)D$; (b) Error vs γ with $r = 9$ and $N_L/(D - N_S) = 0.5$; (c) Error vs N_L with $r = 3$ and $\gamma = 0.01$; (d) Error vs N_L^{-1} with $r = 3$ and $\gamma = 0.01$.

5. EXPERIMENTS

5.1. Sharpness of the Proposed Upper Bound

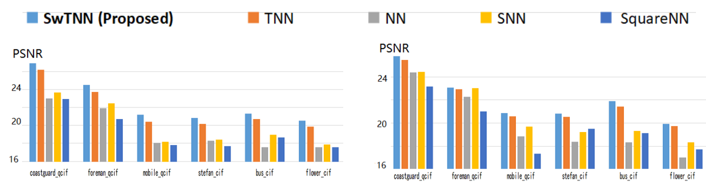
We examine whether the upper bounds in Theorem 1 can predict the scaling behavior of the error. According to Eq. (6), if the upper bound is sharp, then the error $\Delta_{\iota,s}^2/D$ should have the same scaling behavior: approximately linear in the tubal rank r, the corruption ratio γ and the reciprocal of non-corrpted observation number N_L^{-1}. We will check whether the phenomenon occurs using control variable method.

First, we generate the signal tensor $\mathcal{L}^*$ with tubal rank r via $\mathcal{L}^* = \mathcal{P} * \mathcal{Q}$, where $\mathcal{P} \in \mathbb{R}^{d_1 \times r \times d_3}$ and $\mathcal{Q} \in \mathbb{R}^{r \times d_2 \times d_3}$ are sampled from *i.i.d.* standard Gaussian. $\mathcal{L}^*$ is then normalized such that $\|\mathcal{L}^*\|_\infty = 1$. Second, to generate $\mathcal{S}^*$, we first form $\mathcal{S}_0$ with *i.i.d.* uniform distribution $\mathrm{Uni}(0,1)$ and then uniformly select γD elements or γd_2 lateral slices when $\mathcal{S}^*$ represents element sparse errors or slice sparse sample outliers, respectively. Thus the number of corrupted elements $N_S = \gamma D$. Third, we uniformly select N_L elements from the uncorrupted positions of $(\mathcal{L}^* + \mathcal{S}^*)$. Finally, the noise ξ_i are sampled from *i.i.d.* Gaussian $\mathcal{N}(0, \sigma^2)$ with $\sigma = 0.1\|\mathcal{L}^*\|_{\mathrm{F}}/\sqrt{D}$. We consider f-diagonal tensors with $d_1 = d_2 = d = 100$, $d_3 = 30$ and tubal rank $r \in \{3, 6, 9, 12, 15\}$. We choose corruption ratio $\gamma \in \{0.01 : 0.01 : 0.1\}$ and observation ratio $N_L/(D - N_S) \in \{0.4 : 0.1 : 0.9\}$. We solve Problem (4) via ADMM [23]. We set equal weights $w_1 = \cdots = w_{d_3} = 1$. In each setting, the error averaged over 30 runs is reported.

When $\mathcal{S}^*$ represents element sparse errors, Fig. 1 shows the results of the error versus r, γ, N_L and N_L^{-1} by keeping other variables fixed. From sub-figures (a), (b) and (d) in Fig. 1, we can see that the error has approximately linear scaling behavior with respect to r, γ and N_L^{-1}. So, the experimental results are consistent with our expectation in the element sparse error case. Similar phenomena have be found in the cases of slice sparse sample outliers. Since the simulation results are consistent with our expectation, it can be verified that the upper bounds can predict the scaling behavior of the estimation error.

5.2. Effectiveness of the Proposed Estimator

Effectiveness of the SwTNN-based estimator is shown via comparison with other nuclear norm based estimators, i.e., the tensor nuclear norm (SNN) [3], the squared nuclear norm (SquareNN)[24], the tubal nuclear norm (TNN) [16] and the matrix nuclear norm (NN) [25]. *Estimators based on the aforementioned norms are formulated by replacing SwTNN in Problem (4)*, and the corresponding optimization problems are solved by ADMM via our own implementation in Matlab.



(a) $(\rho, \gamma) = (0.5, 0.1)$ (b) $(\rho, \gamma) = (0.8, 0.2)$

Fig. 2: PSNR values for two settings of (ρ, γ) with Gaussian noise level 0.05 on YUV videos.

Given a signal tensor $\mathcal{L}^*$, we consider two settings of the observation ratio ρ and the corruption ratio γ: $(\rho, \gamma) \in \{(0.5, 0.1), (0.8, 0.2)\}$ with Gaussian noise level 0.05. All the involved parameters are manually tuned for better performances unless they have suggested values. The PSNR [22] is used to evaluate recovery quality. For each setting, the experiments are repeated 10 times and the averaged PSNR is reported. We test on six widely used YUV videos[4]: coastguard_qcif, foreman_qcif, mobile_qcif, stefan_cif, bus_cif and flower_cif. Due to the computational limitation, we use the first 32 frames of Y components in each video. That results in three tensors sized $144 \times 176 \times 32$ and three tensors sized $288 \times 352 \times 32$. The PSNR values are reported in Fig. 2. It can be seen that the proposed SwTNN-based estimator has better performances.

6. CONCLUSION

An estimator based on a newly defined SwTNN is proposed for robust low-tubal-rank tensor completion. Non-asymptotic upper bounds on the estimation error are established and further proved to be minimax optimal up to a log factor. Experiments on both synthetic and real datasets demonstrate the sharpness of the proposed upper bounds and the effectiveness of the proposed estimator, respectively.

[4] https://sites.google.com/site/subudhibadri/fewhelpfuldownloads.

References

[1] L. Yuan, Q. Zhao, and J. Cao, "High-order tensor completion for data recovery via sparse tensor-train optimization," in *ICASSP*. IEEE, 2018, pp. 1258–1262.

[2] Q. Zhao, L. Zhang, and A. Cichocki, "Bayesian cp factorization of incomplete tensors with automatic rank determination," *IEEE TPAMI*, vol. 37, no. 9, pp. 1751–1763, 2015.

[3] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE TPAMI*, vol. 35, no. 1, pp. 208–220, 2013.

[4] Q. Zhao, D. Meng, X. Kong, Q. Xie, W. Cao, Y. Wang, and Z. Xu, "A novel sparsity measure for tensor recovery," in *ICCV*, 2015, pp. 271–279.

[5] J. Q. Jiang and M. K Ng, "Exact tensor completion from sparsely corrupted observations via convex optimization," *arXiv preprint arXiv:1708.00601*, 2017.

[6] Q. Zhao, G. Zhou, L. Zhang, A. Cichocki, and S. Amari, "Bayesian robust tensor factorization for incomplete multiway data," *IEEE TNNLS*, vol. 27, no. 4, pp. 736–748, 2016.

[7] W. Hu, D. Tao, W. Zhang, Y. Xie, and Y. Yang, "The twist tensor nuclear norm for video completion," *IEEE TNNLS*, 2016.

[8] M. E Kilmer, K. Braman, N. Hao, and R. C. Hoover, "Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 1, pp. 148–172, 2013.

[9] X-Y. Liu, S. Aeron, V. Aggarwal, et al., "Low-tubal-rank tensor completion using alternating minimization," *arXiv preprint arXiv:1610.01690*, 2016.

[10] X-Y. Liu and X. Wang, "Fourth-order tensors with multidimensional discrete transforms," *arXiv preprint arXiv:1705.01576*, 2017.

[11] Z. Zhang, G. Ely, S. Aeron, N. Hao, and M. Kilmer, "Novel methods for multilinear data completion and denoising based on tensor-svd," in *CVPR*, 2014, pp. 3842–3849.

[12] A. Wang, D. Wei, B. Wang, and Z. Jin, "Noisy low-tubal-rank tensor completion through iterative singular tube thresholding," *IEEE Access*, vol. 6, pp. 35112–35128, 2018.

[13] P. Zhou, C. Lu, Z. Lin, et al., "Tensor factorization for low-rank tensor completion," *IEEE TIP*, vol. 27, no. 3, pp. 1152–1163, March 2018.

[14] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis: Exact recovery of corrupted low-rank tensors via convex optimization," in *CVPR*, 2016, pp. 5249–5257.

[15] P. Zhou and J. Feng, "Outlier-robust tensor pca," in *CVPR*, 2017.

[16] Z. Zhang and S. Aeron, "Exact tensor completion using t-svd," *IEEE TSP*, vol. 65, no. 6, pp. 1511–1526, 2017.

[17] S. Gu, Q. Xie, D. Meng, et al., "Weighted nuclear norm minimization and its applications to low level vision," *IJCV*, vol. 121, no. 2, pp. 183–208, 2017.

[18] O. Klopp, K. Lounici, and A. Tsybakov, "Robust matrix completion," *Probability Theory and Related Fields*, pp. 1–42, 2017.

[19] S. Negahban and M. J. Wainwright, "Estimation of (near) low-rank matrices with noise and high-dimensional scaling," *The Annals of Statistics*, pp. 1069–1097, 2011.

[20] A. Wang and Z. Jin, "Near-optimal noisy low-tubal-rank tensor completion via singular tube thresholding," in *ICDM workshop*, 2017, pp. 553–560.

[21] A. Wang, Z. Lai, and Z. Jin, "Noisy low-tubal-rank tensor completion," *Neurocomputing*, vol. 330, pp. 267–279, 2019.

[22] C. Lu, J. Feng, Z. Lin, et al., "Exact low tubal rank tensor recovery from gaussian measurements," in *IJCAI*, 2018, pp. 1948–1954.

[23] S. Boyd, N. Parikh, et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[24] C. Mu, B. Huang, J. Wright, and D. Goldfarb, "Square deal: Lower bounds and improved relaxations for tensor recovery," in *ICML*, 2014, pp. 73–81.

[25] E. J. Candès and T. Tao, "The power of convex relaxation: near-optimal matrix completion," *IEEE TIT*, vol. 56, no. 5, pp. 2053–2080, 2010.