From Simulated to Visual Data: A Robust Low-Rank Tensor Completion Approach Using l_p-Regression for Outlier Resistance

Qi Liu, *Member, IEEE*, Xiaopeng Li, *Member, IEEE*, Hui Cao, *Member, IEEE*, and Yuntao Wu, *Senior Member, IEEE*

Abstract— Low-rank tensor completion (LRTC) that aims to restore the latent clean data from an incomplete and/or degraded observation, shows promising results in ubiquitous tensorial data completion applications. Most tensor completion approaches are vulnerable to outliers since their derivations are based on ℓ_2-space to be robust against Gaussian noise. In this work, to tackle this issue, ℓ_p-regression ($0 < p < 2$) is employed to achieve outlier resistance, where a factored form of tensor train (TT)-format representation is regularized by the low-TT-rank prior to exploit the inter-fibers correlation. On the basis of that, an effective iterative ℓ_p-regression TT completion method (referred to ℓ_p-TTC) is proposed, with the advantage of not requiring the hard-to-determine user-defined weights in TT rank model. Extensive experiment results are presented to demonstrate the outlier resistance of the proposed ℓ_p-TTC, and showing the effective and superior performance in both bistatic MIMO radar localization and color image inpainting and denoising, compared with state-of-the-art tensor completion approaches.

Index Terms— Low-rank tensor completion (LRTC), alternating minimization, ℓ_p-regression, tensor train (TT) rank, multiple-input multiple-output (MIMO) radar, color image inpainting and denoising.

I. INTRODUCTION

TENSORS, a.k.a., multidimensional multi-way arrays, are an effective higher-order generation of vectors and matrices [1]–[3]. The big data era endows immense opportunities for ubiquitous tensors and their decompositions applications by leveraging big data generated from widespread sensors and ever-growing computing capability. Benefiting from the power of multi-linear algebra as their mathematical backbone [4]–[6],

they enable more efficient paradigm shift toward models for versatile data analysis tools, compared to their counterparts with standard flat-view matrix models. Therefore, they have become increasingly popular and omnipresent across diverse fields, ranging from signal and image processing, graphics, computer vision to data mining, machine learning, and brain modeling [7]–[11].

The success of low-rank matrix completion (LRMC) spurs numerous researchers to extend its concept to low-rank tensor completion (LRTC). In fact, due to the imperfection of data collection, transmission and compression, the high-dimensional data of interest are usually incomplete, hence, data completion is fertile research ground to ensure data integrity that merits further investigation. To this end, LRTC acts as a powerful computational tool for extracting valuable information from tensorial data. Given an observed incomplete high-dimensional data, the objective of LRTC aims to predict plausible missing entries by exploiting their correlation with the observed ones. This problem has been extensively studied in literature, and a growing number of promising LRTC methods have been proposed, instances include Bayesian CANDE-COMP/PARAFAC (Bayesian CP) [12], high accuracy LRTC (HaLRTC) [8], fast LRTC (FaLRTC) [8], simple LRTC via tensor train (SiLRTC-TT) [7], tensor completion by parallel matrix factorization via tensor train (TMac-TT) [7], fixed-point iterative LRTC (FP-LRTC) [13], and tensor singular value decomposition (t-SVD) [14]. In [12], CP tensor factorization technique is introduced in tensor completion, where the latent CP rank is automatically determined by exploiting a fully Bayesian treatment. Still, however, it is challenging for LRTC because of the NP-hard CP rank minimization. Therefore, Tucker rank, as a promising alternative, is defined with the ranks of unfolded matrices along each order of a tensor. On the basis of Tucker decomposition, many works are proposed for image completion, and instances include [8], [15]–[18]. Liu *et al*. [8] proposes a pioneer idea to approximate the hard-to-handle CP rank function by the Tucker-based tensor nuclear norm, with resulting two methods, namely, HaLRTC and FaLRTC, by combining nuclear norms of all matrices unfolded along each mode. Both of them achieve satisfactory performance, while ignoring the structural information of the tensor data. In [15], Tucker decomposition with the

alternating least squares algorithm and an additional lowpass Gaussian filtration are exploited to enhance the quality of image inpainting. Reference [16] adopts a doubly weighted strategy for nuclear norm along each mode to characterize global sparsity prior of tensor. In Tucker rank, we notice that unfolded matrices have unbalanced sizes, namely, varying numbers between rows and columns. Since the rank of an unfolded matrix is bounded by the minimal number of rows and columns, the inefficiency of rank minimization due to unbalance would render the resulting algorithms for capturing the global information of a tensor in vain.

To address that, tensor train (TT) rank [19] is defined, which consists of ranks of matrices unfolded in a well-balanced matricization way, i.e., unfolding the tensor along permutations of modes. On the basis of that, Bengua *et al.* [7] introduces the SiLRTC-TT and TMac-TT approaches, which are extended from the SiLRTC [8] and multi-linear matrix factorization model, respectively. Using the matrix product state representation of TT decomposition, a tensor completion algorithm by alternating minimization under the TT model (TCAM-TT) is proposed in [20], which presents noticeably superior performance for a variety of real settings. On the other hand, to incorporate the data factor priors and tensor multi-linear structural information, a simultaneous tensor decomposition and completion (STDC) method is developed in [21]. Similar with the LRMC using approximate singular value decomposition (SVD) based fixed-point continuation algorithm (FPCA) [22], the FP-LRTC method in [13] is proposed by employing operator splitting and convex relaxation techniques. To characterize informational and structural complexity of multi-linear data, t-SVD [14] is employed with applications to three-dimensional (3-D) and 4-D video data completion and denoising from limited samples. Nevertheless, the tensor nuclear norm minimization (TNNM)-based algorithms suffer from high computation cost of multiple SVDs at each iteration. What is more, how to determine the best user-defined weights in TT rank is still an open problem. As mentioned above, most of the existing TNNM algorithms are based on the Frobenius norm to robustify against the Gaussian noise.

Whereas non-Gaussian noises or outliers commonly exist in many different types of data, e.g., image, video, text and bioinformatics. Many existing LRTC-based approaches explicitly or implicitly assume that the noise is Gaussian distributed [23], [24]. Therefore, the performance of these approaches may severely degrade due to the existence of outliers. Moreover, tensor decomposition with low-rank structures in TNNM is sensitive to the presence of outliers, as it attempts to force the outliers to fit the low-rank structure. It results in that the conventional TNNM techniques fail to work properly when the observations contain outliers as their derivations are based on the ℓ_2-space, e.g., Frobenius norm of tensor residual error. Herein, outliers refer to outlying entries whose values are abnormally large. Numerous approaches have employed the fact that the entry-wise ℓ_p-norm with $p < 2$ is less sensitive to outliers than the Frobenius norm for robust LRTC. In [25] and [26], ℓ_p-norm ($0 < p < 2$) is involved to assist to address additive impulsive noise, with the resulting iteratively

reweighted tensor SVD (IR-t-SVD) and ℓ_p-PARAFAC methods, to recover the target tensor, respectively. This is also under out consideration for LRTC purposes.

In this work, we propose to complete tensorial data from both randomly missing entries and outliers based on a factored form of TT-format representation. In our model, the latent tensor is regularized by the low-TT-rank prior to exploit the inter-fibers correlation, which is motivated by the fact that TT rank enables better to capture the global correlation between tensor entries for LRTC compared to its counterpart, namely, Tucker rank [7], [27]. We derive a promising method using the combination of alternating minimization and alternating direction method of multipliers (ADMM) under the ℓ_p-regression ($0 < p < 2$) framework. The main contributions of this work are summarized as follows:

- Owing to the fact that it is challenging to determine the best user-defined weights in TT rank model, we propose to employ ℓ_p-regression based on tensor factorization for solving LRTC problems.
- To be robust against both missing entries and outliers, an effective iterative ℓ_p-regression tensor train completion method (ℓ_p-TTC) is developed, where ADMM is applied as the solver under the framework of augmented Lagrangian multiplier.
- Finally, compared with state-of-the-art tensor completion approaches, numerical experiments have been conducted to verify the efficacy and superiority of the proposed ℓ_p-TTC on both simulated and visual data, for bistatic multiple-input multiple-output (MIMO) radar localization and color image inpainting and denoising, respectively. The code is available at https://sites.google.com/site/qiliucityu/discussion.

The rest of the paper is organized as follows. In Section II, the LRTC problem is formulated, where notations and preliminaries on tensor decomposition are introduced. In Section III, we propose the iterative ℓ_p-regression tensor completion method based on the TT model. Numerical experiment results are presented on LRTC with applications to bistatic MIMO radar localization, and color image inpainting and denoising in Section IV. Finally, conclusions are drawn in Section V.

II. BACKGROUND

A. Preliminaries

In this work, we adopt the tensor notations and preliminaries from [5]. The order or mode of a tensor is defined as the number of tensor dimensions. Scalars, a.k.a., zero-order tensors, are denoted by lowercase letters ($x, y, z, \cdots$). The first-order and second-order tenosrs, namely, vectors and matrices, are denoted by boldface lowercase letters ($\mathbf{x}, \mathbf{y}, \mathbf{z}, \cdots$) and uppercase boldface letters ($\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \cdots$), respectively. Tensors with higher order are represented by calligraphic uppercase boldface letters ($\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \cdots$). An N-mode or N-order tensor, is denoted as $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$, where I_k is the dimension along mode k and its element is expressed by $\mathcal{X}(i_1, \cdots, i_k, \cdots, i_N)$, $1 \leq i_k \leq I_k$, $1 \leq k \leq N$. A mode-k fiber of a tensor $\mathcal{X}$ is a vector defined by fixing all indices except for the kth index.

Definition 1 (Mode-k Matricization or Unfolding): A tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ is unfolded or reshaped into a matrix
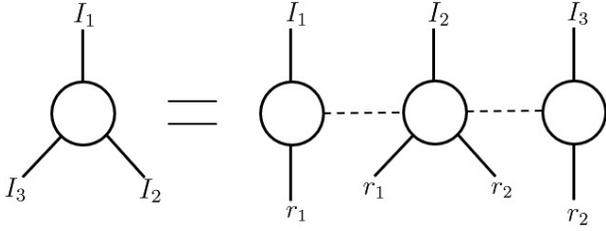
Fig. 1. Illustration of the TT-format.

$\mathbf{X}_{(k)} \in \mathbb{R}^{I_k \times (I_1 \cdots I_{k-1} I_{k+1} \cdots I_N)}$ by arranging the mode-k fibers to be the columns of the resulting matrix. The corresponding mapping relation between $\mathcal{X}$ and $\mathbf{X}_{(k)}$ is:

$$\mathbf{X}_{(k)}(i_k, j) = \mathcal{X}(i_1, \cdots, i_k, \cdots, i_N),$$
$$j = 1 + \sum_{n=1, n \neq k}^{N} (i_n - 1) J_n \quad (1)$$

where $J_n = \prod_{m=1, m \neq k}^{n-1} I_m$.

Definition 2 (Tensor Connect Product [19], [28]): For two 3rd-order tensors $\mathcal{X} \in \mathbb{R}^{r_0 \times I_1 \times r_1}$ and $\mathcal{Y} \in \mathbb{R}^{r_1 \times I_2 \times r_2}$, the tensor connect product is defined as:

$$\mathcal{X}\mathcal{Y} = \text{reshape}((\mathcal{X}_{(3)})^{\text{T}} \times (\mathcal{Y}_{(1)})) \in \mathbb{R}^{r_0 \times (I_1 I_2) \times r_2} \quad (2)$$

where reshape($\cdot$) refers to a reshape operator which returns an array of specified dimensions with the same entries as the input data. Tensor connect product is to calculate the product of two 3rd-tensors where the tensors are first unfolded into matrices, and then the product of two matrices is computed. Notice that the last dimension of the former must be equal to the first dimension of the latter.

Definition 3 (Tensor Permutation From Lemma 2 [20]): Given Nth-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$, the kth TT permutation is $\mathcal{X}^{\mathcal{P}_k} \in \mathbb{R}^{I_k \times I_{k+1} \times \cdots \times I_N \times I_1 \cdots I_{k-1}}$ with element of $\mathcal{X}^{\mathcal{P}_k}(j_k, \cdots, j_N, j_1, \cdots, j_{k-1}) = \mathcal{X}(j_1, \cdots, j_N)$, for any $j_k \in [1, I_k]$. Therefore, we get $\mathcal{X}^{\mathcal{P}_k} = Trace(\mathcal{X}_k \mathcal{X}_{k+1} \cdots \mathcal{X}_N \mathcal{X}_1 \cdots \mathcal{X}_{k-1})$. As shown in Fig. 1, the TT-format, a.k.a., TT network, is illustrated [29], [30].

B. Problem Formulation

Given a low-rank incomplete tensor $\mathcal{T}_\Omega \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ with missing entries where Ω is a subset of the complete set of entries $[I_1] \times \cdots \times [I_N]$, with $[I]$ being the list $\{1, \cdots, I\}$. The set Ω contains all tensor coordinates corresponding to the observed entries in tensor $\mathcal{T}$, and the subscript $(\cdot)_\Omega$ denotes the projection on the known entries. The task of LRTC is to complete the Nth-order tensor $\mathcal{T}$ from its known entries given by the index set Ω. Mathematically, it is formulated as a rank minimization problem by incorporating the low-rank information:

$$\min_{\mathcal{X}} \text{rank}(\mathcal{X}), \quad \text{s.t. } \mathcal{X}_\Omega = \mathcal{T}_\Omega. \quad (3)$$

Due to the combinatorial nature of the rank function, the rank minimization (3), however, is NP-hard. Many attempts have been proposed to relax the underlying tensor rank, including CP rank, Tucker rank, and TT rank.

For the Tucker rank, (3) is written by the following Tucker rank optimization problem [31], [32]:

$$\min_{\mathbf{X}_{(k)}} \sum_{k=1}^{N} \alpha_k \text{rank}(\mathbf{X}_{(k)}), \quad \text{s.t. } \mathcal{X}_\Omega = \mathcal{T}_\Omega. \quad (4)$$

where $\{\alpha_k\}_{k=1}^{N}$ are defined as the nonnegative weights for rank$(\mathbf{X}_{(k)})$ satisfying the condition $\sum_{k=1}^{N} \alpha_k = 1$. Nevertheless, Tucker rank is conceptually limited by the small upper bound of each individual rank and may not be suitable for describing global information of tensor [7], because of its ranks of matrices stemming from unbalanced matricization scheme, which is only efficient for more balanced matrices.

To deal with that, another type of tensor rank is the TT rank. For the ease of optimization, a popular and practical solution is to replace the nonconvex rank by convex nuclear norm, resulting in the TT nuclear norm optimization problem [7], [33]:

$$\min_{\mathcal{X}} \sum_{k=1}^{N} \alpha_k ||\mathbf{X}_{(k)}||_*, \quad \text{s.t. } \mathcal{X}_\Omega = \mathcal{T}_\Omega \quad (5)$$

with TT rank combining more balanced unfolded matrices compared to Tucker rank, for tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$. Herein, $\mathbf{X}_{(k)}$ are obtained by matricizing along mode k, $k = 1, \cdots, N$. The nuclear norm $|| \cdot ||_*$ represents the sum of singular values λ_k, namely, $||\mathbf{X}||_* = \sum_{k=1}^{\min(m,n)} \lambda_k$ for any $\mathbf{X} \in \mathbb{R}^{m \times n}$.

Corrupted by the noise, (5) is modified as:

$$\min_{\mathcal{X}} \sum_{k=1}^{N} \alpha_k ||\mathbf{X}_{(k)}||_*, \quad \text{s.t. } ||\mathcal{X}_\Omega - \mathcal{T}_\Omega||_F \leq \delta \quad (6)$$

where $\delta > 0$ stands for the noise tolerance that controls the fitting error, and $|| \cdot ||_F$ denotes the Frobenius norm of a tensor by $||\mathcal{X}||_F = ||vec(\mathcal{X})||_2$. $|| \cdot ||_2$ and $vec(\cdot)$ denote the ℓ_2-norm and vectorization operators, respectively. Although (6) works well in the presence of additive Gaussian disturbance, its performance can significantly degrade when $\mathcal{T}_\Omega$ contains outliers. Moreover, how to choose the best user-defined parameters $\{\alpha_k\}_{k=1}^{N}$ is still an open problem.

III. ITERATIVE ℓ_p-REGRESSION TT COMPLETION

It is well known that ℓ_2-space optimization cannot resist impulsive noise effectively. In contrast, ℓ_p-norm enables to be robust against impulsive noise as it reduces the effect of outliers via calculating the residual to power of p with $0 < p < 2$. Therefore, in this work, we recast the problem (6) using ℓ_p-regression:

$$\min_{\mathcal{X}} ||\mathcal{X}_\Omega - \mathcal{T}_\Omega||_p^p \quad (7)$$

to achieve outlier resistance. Obviously, (7) is hard to solve because of the high nonconvexity of ℓ_p-norm.

Motivated by the Definition 2, (7) is written as:

$$\min_{\mathcal{X}_1, \mathcal{X}_2, \cdots, \mathcal{X}_N} ||(\mathcal{X}_1 \cdot \mathcal{X}_2 \cdots \mathcal{X}_N)_\Omega - \mathcal{T}_\Omega||_p^p \quad (8)$$

based on the assumption of TT rank information.

Remark 1: Regarding the TT rank of a sequence of tensors, they can be determined directly by the diagonal singular

matrices $\{\boldsymbol{\lambda}^{[k]}\}_{k=1}^N \in \mathbb{R}^{r_k \times r_k}$. Based on the Vidal's decomposition [34], the tensor can be rewritten by

$$\mathcal{X}_k = \sum_{i_1, \cdots, i_N} \mathbf{u}^{[1\cdots k]i_1 \cdots i_k} \boldsymbol{\lambda}^{[k]} \mathbf{v}^{[k+1\cdots N]i_{k+1} \cdots i_N} \quad (9)$$

where $\mathbf{u}^{[1\cdots k]i_1 \cdots i_k} = \boldsymbol{\Gamma}_{i_1}^{[1]} \boldsymbol{\lambda}^{[1]} \cdots \boldsymbol{\Gamma}_{i_k}^{[k]} \otimes_{\ell=1}^k \mathbf{e}_{i_\ell}$ and $\mathbf{v}^{[k+1\cdots N]i_{k+1} \cdots i_N} = \boldsymbol{\Gamma}_{i_{k+1}}^{[k+1]} \boldsymbol{\lambda}^{[k+1]} \cdots \boldsymbol{\Gamma}_{i_N}^{[N]} \otimes_{\ell=k+1}^N \mathbf{e}_{i_\ell}$.

Moreover, borrowing the results from the TT decomposition [19], i.e., tensor factorization, the element of tensor $\mathcal{X}$ can be expressed as:

$$\mathcal{X}(i_1, \cdots, i_k, \cdots, i_N)$$
$$= \mathcal{X}_1(:, i_1, :) \cdots \mathcal{X}_k(:, i_k, :) \cdots \mathcal{X}_N(:, i_N, :) \quad (10)$$

where $\mathcal{X}(:, i_k, :) \in \mathbb{R}^{r_{k-1} \times I_k \times r_k}$, $k = 1, \cdots N$, with r_k being the TT rank. It is noteworthy that $r_0 = r_N = 1$. Specifically, $\mathcal{X}_1$ and $\mathcal{X}_N$ are actually the boundary matrices with the sizes of $r_0 \times I_1 \times r_1 (r_0 = 1)$ and $r_{N-1} \times I_N \times r_n (r_n = 1)$, respectively.

From (10), we observe that

$$\mathcal{X}(i_1, \cdots, i_N) = Trace(\mathcal{X}_1(:, i_1, :) \cdots \mathcal{X}_N(:, i_N, :)) \quad (11)$$

for the scalar $\mathcal{X}(i_1, \cdots, i_N)$, where $Trace(\cdot)$ denotes the trace operator. Hence, (8) can be further optimized by involving the trace operator, that is:

$$\min_{\mathcal{X}_1, \mathcal{X}_2, \cdots, \mathcal{X}_N} ||(Trace(\mathcal{X}_1 \cdot \mathcal{X}_2 \cdots \mathcal{X}_N))_\Omega - \mathcal{T}_\Omega||_p^p. \quad (12)$$

Different with [2], [35], tensor factorization is exploited to avoid SVD computation. After determining each tensor $\{\mathcal{X}_k\}_{k=1}^N$, the target tensor is obtained as $\mathcal{X} = \mathcal{X}_1 \cdot \mathcal{X}_2 \cdots \mathcal{X}_N$. In this work, we adopt the alternating minimization strategy to solve (12), we obtain:

$$\begin{cases} \mathcal{X}_1^{t+1} = \arg \min_{\mathcal{X}_1} ||(Trace(\mathcal{X}_1 \cdot \mathcal{X}_2^t \cdots \mathcal{X}_N^t))_\Omega - \mathcal{T}_\Omega||_p^p \\ \mathcal{X}_2^{t+1} = \arg \min_{\mathcal{X}_2} ||(Trace(\mathcal{X}_1^{t+1} \cdot \mathcal{X}_2 \cdots \mathcal{X}_N^t))_\Omega - \mathcal{T}_\Omega||_p^p \\ \quad \vdots \\ \mathcal{X}_N^{t+1} = \arg \min_{\mathcal{X}_N} ||(Trace(\mathcal{X}_1^{t+1} \cdot \mathcal{X}_2^{t+1} \cdots \mathcal{X}_N))_\Omega - \mathcal{T}_\Omega||_p^p \end{cases}$$
$$(13)$$

with updating one tensor $\mathcal{X}_k$ by fixing others for each time step, $1 \leq k \leq N$. Now, we focus on solving (13) for an $\mathcal{X}_k$:

$$\mathcal{X}_k^{t+1}$$
$$= \arg \min_{\mathcal{X}_k} \left| \left| (Trace(\mathcal{X}_1^{t+1} \cdots \mathcal{X}_k \cdots \mathcal{X}_N^t))_\Omega - \mathcal{T}_\Omega \right| \right|_p^p$$
$$= \arg \min_{\mathcal{X}_k} \left| \left| (Trace(\mathcal{X}_k \mathcal{X}_{k+1}^t \cdots \mathcal{X}_N^t \mathcal{X}_1^{t+1} \cdots \mathcal{X}_{k-1}^{t+1}))_\Omega^{\mathcal{P}_k} - \mathcal{T}_\Omega \right| \right|_p^p$$
$$= \arg \min_{\mathcal{X}_k} \left| \left| (Trace(\mathcal{X}_k \mathcal{U}))_\Omega^{\mathcal{P}_k} - \mathcal{T}_\Omega \right| \right|_p^p \quad (14)$$

where $(\cdot)^{\mathcal{P}_k}$ denotes the tensor permutation, and $\mathcal{U} = \mathcal{X}_{k+1}^t \cdots \mathcal{X}_N^t \mathcal{X}_1^{t+1} \cdots \mathcal{X}_{k-1}^{t+1} \in \mathbb{R}^{r_k \times (I_{k+1} \cdots I_N I_1 \cdots I_{k-1}) \times r_{k-1}}$. The second equality holds due to the Definition 3. According to Definition 1, as the TT-format unfolded matrices $\mathbf{X}_{(k)}$ share the same entries of the tensor $\mathcal{X}$, (14) is equivalently transformed into:

$$\mathcal{X}_k^{t+1} = \arg \min_{\mathcal{X}_{(k)}} \left| \left| \left((Trace(\mathcal{X}_k \mathcal{U}))_\Omega^{\mathcal{P}_k} \right)_{(k)} - (\mathcal{T}_\Omega)_{(k)} \right| \right|_p^p \quad (15)$$

by using mode-k matricization. Recall that $\mathbf{X}_{(k)}$ is obtained by unfolding along k modes and thus its rank captures the correlation between k modes and the other $N - k$ modes. Hence, it enables better to capture the global correlation of a tensor in (15) as it encompasses of correlations between permutations of all modes [7]. On the basis of that each slice of $\mathcal{X}_k$, namely, $\mathcal{X}_k(:, i_k, :)$, $1 < i_k < I_k$, corresponds to each row of $\mathbf{X}_{(k)}$, we can solve (15) by separate manner with I_k similar subproblems. Regarding i_kth subproblem, we have:

$$\mathcal{X}_k^{t+1}(:, i_k, :)$$
$$= \arg \min_{\mathcal{X}_{(k)}} \left| \left| \left((Trace(\mathcal{X}_k \mathcal{U})_\Omega^{\mathcal{P}_k} \right)_{(k)} (i_k, :) \right. \right.$$
$$\left. \left. - (\mathcal{T}_\Omega)_{(k)} (i_k, :) \right| \right|_p^p$$
$$= \arg \min_{\mathbf{X}} \sum_{j \in \Omega/i_k} \left| \left| Trace \left(\mathbf{X} \times (\mathcal{U}_{\Omega/i_k}^{\mathcal{P}_k})_{(k)}(:, j, :) \right) \right. \right.$$
$$\left. \left. - (\mathcal{T}_\Omega)_{(k)} (i_k, j) \right| \right|_p^p \quad (16)$$

where $\mathbf{X} \in \mathbb{R}^{r_{k-1} \times r_k}$ and Ω/i_k denotes a subset of the complete set of entries $[r_k] \times [I_{k+1} \cdots I_N I_1 \cdots I_{k-1}] \times [r_{k-1}]$ excluding i_k, $i_k = 1, \cdots, I_k$. The second equality is derived due to the entry-wise summation of all entries. For each j, $(\mathcal{U}_{\Omega/i_k}^{\mathcal{P}_k})_{(k)}$ $(:, j, :) \in \mathbb{R}^{r_k \times 1 \times r_{k-1}}$ can be regarded as a matrix of size $r_k \times r_{k-1}$. Following the results of linear algebra, $Trace(\mathbf{U}\mathbf{V}) = vec(\mathbf{V}^T)^T vec(\mathbf{U})$ for any matrices $\mathbf{U}$ and $\mathbf{V}$. Therefore, (16) is further modified as:

$$\mathcal{X}_k^{t+1}(:, i_k, :) = \arg \min_{\mathbf{X}} \sum_{j \in \Omega/i_k} \left| \left| vec \left((\mathcal{U}_{\Omega/i_k}^{\mathcal{P}_k})_{(k)}(:, j, :)^T \right)^T \right. \right.$$
$$\left. \left. \times vec(\mathbf{X}) - (\mathcal{T}_\Omega)_{(k)} (i_k, j) \right| \right|_p^p \quad (17)$$

where $(\cdot)^T$ stand for the transpose operator. For the ease of simplicity, (17) is rewritten as:

$$\mathcal{X}_k^{t+1}(:, i_k, :) = \arg \min_{\mathbf{X}} ||\mathbf{A} vec(\mathbf{X}) - \mathbf{y}||_p^p \quad (18)$$

where $\mathbf{A} \in \mathbb{R}^{[j \in \Omega/i_k] \times (r_{k-1} r_k)}$ and $\mathbf{y} \in \mathbb{R}^{[j \in \Omega/i_k]}$. Let $\mathbf{e}$ equal to $\mathbf{A} vec(\mathbf{X}) - \mathbf{y}$, then it is necessary to convert the unconstrained problem into a constrained one, that is:

$$\min_{\mathbf{e}} ||\mathbf{e}||_p^p, \quad \text{s.t. } \mathbf{e} = \mathbf{A} vec(\mathbf{X}) - \mathbf{y} \quad (19)$$

The augmented Lagrangian of (19) is:

$$\mathcal{L}_\mu(\mathbf{e}, \mathbf{X}, \Lambda) = ||\mathbf{e}||_p^p + \langle \Lambda, \mathbf{A} vec(\mathbf{X}) - \mathbf{y} - \mathbf{e} \rangle$$
$$+ \frac{\mu}{2} ||\mathbf{A} vec(\mathbf{X}) - \mathbf{y} - \mathbf{e}||_2^2 \quad (20)$$

where Λ is a vector of Lagrange multipliers. It is difficult to minimize the objective function over multiple sets of variables. The popular and practical solution is applying ADMM [36] as the solver to alternatively optimize one set of variables with others fixing. To solve (20), the main steps corresponding to ADMM are concluded as follows:

$$\mathbf{X}^{l+1} = \arg \min_{\mathbf{X}} \mathcal{L}_\mu(\mathbf{e}^l, \mathbf{X}, \Lambda^l) \quad (21)$$

$$\mathbf{e}^{l+1} = \arg \min_{\mathbf{e}} \mathcal{L}_\mu(\mathbf{e}, \mathbf{X}^{l+1}, \Lambda^l) \quad (22)$$

$$\Lambda^{l+1} = \Lambda^l + \mu(\mathbf{A} vec(\mathbf{X}^{l+1}) - \mathbf{y} - \mathbf{e}^{l+1}). \quad (23)$$

Algorithm 1: ℓ_p-TTC

Require: $\mathcal{T}_\Omega \in \mathbb{R}^{I_1 \times \cdots \times I_N}$, Ω and $\mathbf{r} = [r_1, r_2, \cdots, r_N]$, p, $\epsilon = 10^{-6}$, $\xi = 1.1$.

  Initialize: Randomize $\mathcal{X}_i$

  for $t = 1, 2, \cdots$ **do**

    for $k = 1, 2, \cdots, N$ **do**

$$\mathcal{X}_k^{t+1} = \arg\min_{\mathcal{X}_{(k)}} \left\| \left(Trace(\mathcal{X}_k \mathcal{U})_\Omega^{\mathcal{P}_k}\right)_{(k)} - (\mathcal{T}_\Omega)_{(k)} \right\|_p^p.$$

      for $i_k = 1, 2, \cdots, I_k$ **do**

$$\mathcal{X}_k^{t+1}(:, i_k, :) = \arg\min_{\mathbf{X}} \|\mathbf{A}vec(\mathbf{X}) - \mathbf{y}\|_p^p.$$

      1) Convert the unconstrained problem into the constrained one, namely:

$$\min_{\mathbf{e}} \|\mathbf{e}\|_p^p, \quad \text{s.t. } \mathbf{e} = \mathbf{A}vec(\mathbf{X}) - \mathbf{y}.$$

      2) Compute the corresponding augmented Lagrangian, then ADMM is applied as the solver.

      for $l = 1, 2, \cdots$ **do**

        i. Update $\mathbf{X}^{l+1}$ via (24)

        ii. Update $\mathbf{e}^{l+1}$ via (25)

          Case 1: $p = 1$, solved by (26);

          Case 2: $0 < p < 1$, solved by (27);

          Case 3: $1 < p < 2$, solved by (30).

        iii. Update Λ^{l+1} via (23)

      end for

      end for

    end for

    Stop if the stopping criterion is satisfied.

  end for

Ensure: Target tensor $Trace(\mathcal{X}_1^{t+1} \mathcal{X}_2^{t+1} \cdots \mathcal{X}_N^{t+1})$

Step 1: Update $\mathbf{X}^{l+1}$

In subproblem (21), it can be equivalently converted into a ℓ_2-norm minimization problem:

$$\arg\min_{\mathbf{X}} \left\| \mathbf{A}vec(\mathbf{X}) - (\mathbf{e}^l - \frac{\Lambda^l}{\mu} + \mathbf{y}) \right\|_2^2 \tag{24}$$

which has a closed-form solution $\mathbf{X} = reshape(\mathbf{A}^{-1}(\mathbf{e}^l - \frac{\Lambda^l}{\mu} + \mathbf{y}), [r_{k-1}, r_k])$, and $reshape(\mathbf{x})$ denotes to reshape a vector $\mathbf{x} \in \mathbb{R}^{r_{k-1}r_k}$ into a matrix $\mathbf{X} \in \mathbb{R}^{r_{k-1} \times r_k}$.

Step 2: Update $\mathbf{e}^{l+1}$

In the second subproblem (22), we actually solve the following problem:

$$\min_{\mathbf{e}} \frac{1}{\mu} \|\mathbf{e}\|_p^p + \frac{1}{2} \|\mathbf{e} - \mathbf{z}\|_2^2, \quad 0 < p < 2 \tag{25}$$

where $\mathbf{z} = \mathbf{A}vec(\mathbf{X}) + \frac{\Lambda^l}{\mu} - \mathbf{y}$).

For the case of $p = 1$, based on the soft-thresholding operator [37], we can obtain the closed-form solution as:

$$e_i = \text{sign}(z_i) \max(|z_i| - 1/\mu, 0) \tag{26}$$

where e_i and z_i are the ith entry of $\mathbf{e}$ and $\mathbf{z}$, respectively.

For the case of $0 < p < 1$, we follow the result in [38] to get the corresponding solution:

$$e_i = \begin{cases} 0, & \text{if } |z_i| \le h_a \\ \text{sign}(z_i)\beta_\star, & \text{if } |z_i| > h_a \end{cases} \tag{27}$$

where $h_a = \beta_a + \frac{p}{\mu}\beta_a^{p-1}$ with $\beta_a = (\frac{2}{\mu}(1-p))^{\frac{1}{2-p}}$ and $\beta_\star \in (\beta_a, |z_i|)$ is obtained by

$$\beta + \frac{1}{\mu} p\beta^{p-1} = |z_i|. \tag{28}$$

The solution of (28) is the convergent point of the following iteration:

$$\beta_{l+1} = \rho(\beta_l) \tag{29}$$

where $\rho(\beta) = |z_i| - \frac{1}{\mu} p\beta^{p-1}$ with the initial value of $\beta_\star \in (\beta_a, |z_i|)$.

For the case of $1 < p < 2$, the problem (25) is convex, and can be solved by [39]:

$$e_i = \begin{cases} \arg\min\{g(0), g(r_i^+)\}, & \text{if } z_i \ge 0 \\ \arg\min\{g(0), g(r_i^-)\}, & \text{if } z_i < 0 \end{cases} \tag{30}$$

where $g(e_i) = \frac{1}{2}(e_i - z_i)^2 + \frac{1}{\mu}|e_i|^p$, r_i^+ and r_i^- are the solutions of $g'(e_i)$ in the case of $z_i \ge 0$ and $z_i < 0$, respectively. For $z_i \ge 0$, $g'(e_i)$ is monotonically creased in $[0, z_i]$ and $g'(0)g'(z_i) < 0$. Thereby, r_i^+ locates in $[0, z_i]$, which can be found by the bisection method with a complexity of $\mathcal{O}(1)$. In the same way, r_i^- can be got from solving $g'(e_i)$ via the bisection method in the feasible region $[z_i, 0]$.

After obtaining the $\mathbf{X}^{l+1}$ and $\mathbf{e}^{l+1}$, the Lagrange multiplier matrix Λ^{l+1} is updated as (23). The stepsize μ is computed by $\mu^{l+1} = \xi\mu^l$, where $\xi > 1$ is a constant. The pseudocode of the proposed ℓ_p-TTC is provided in Algorithm 1. The convergence criterion of the proposed ℓ_p-TTC is characterised by

$$\frac{\|\mathcal{X}^{t+1} - \mathcal{X}^t\|_F}{\|\mathcal{X}^t\|_F} \le \epsilon \tag{31}$$

with the relative error of estimate between two successive iterations. In this work, ϵ is set at 10^{-6}. Although the convergence theory of ADMM has been well-established for a variety of ADMM variants in [36], including two-block ADMM, they cannot be directly applicable for our problem with missing entries. The reasons are twofold: 1) ℓ_p-regression is utilized to minimize the fitting error instead of appearing as a regularization term to promote sparsity, and ℓ_p-norm with $0 < p < 1$ is highly nonconvex; 2) Because of the TT model, a sequence of tensor connect products exist, which makes the resulting ℓ_p-regression optimization problem nonconvex. Thereby, it is still an open problem that deserves more studies in the future. However, the proposed method is indeed convergent,[1] and verified by the empirical results in Figs. 2 and 3. From the simulation results, we can see that the proposed method enjoys not only smaller MSE, but also faster convergence in terms of very few iterations, less than 5 times.

In this work, the proposed ℓ_p-TTC can deal with the ℓ_p-regression tensor completion problem at $0 < p < 2$. In the presence of both missing entries and outliers corrupted observations, the proposed method can achieve the best performance at $p = 1$, as shown in Fig. 4. Taking the example of color image inpainting and denoising, we can clearly and

---

[1]The proposed method in the case of $p < 1$ only converges to a local point because of the non-convexity of ℓ_p-norm.
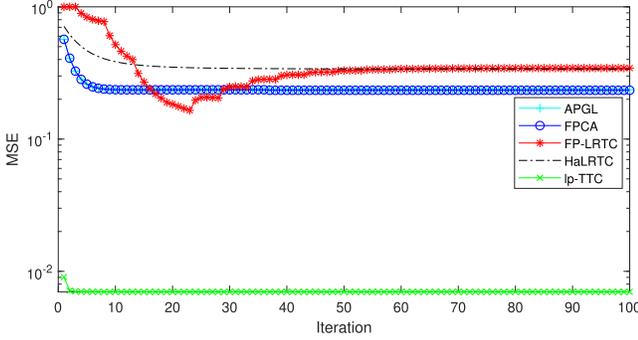
Fig. 2. Illustration of the convergence of the proposed method, as compared to different algorithms, where is examined on the bistatic MIMO radar tensorial data, and the parameter settings are the same in Section IV.
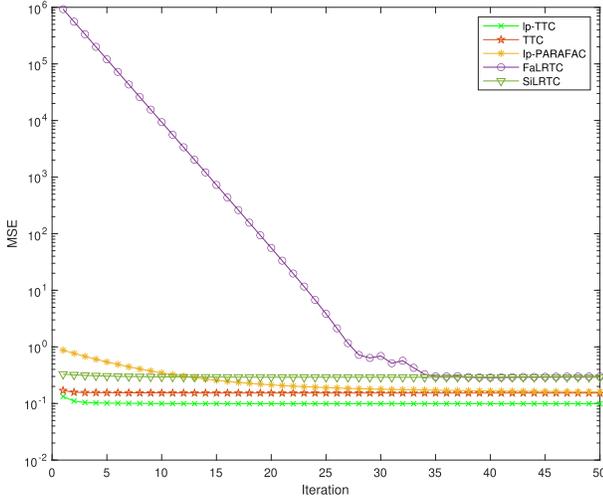

Fig. 3. Illustration of the convergence of the proposed method, as compared to different algorithms, where is examined on the color image inpainting and denoising tensorial data, and the test image is *Windows*.
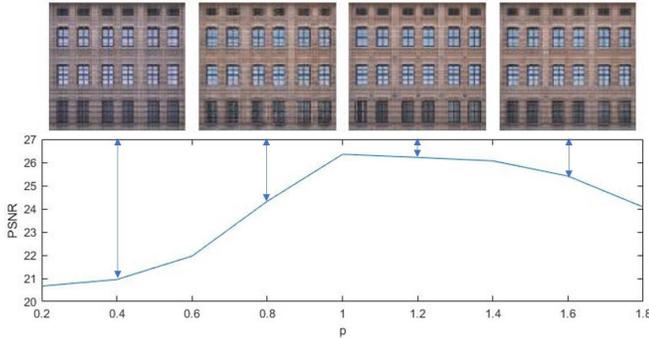

Fig. 4. Illustration of PSNR conditioned on different p, $0 < p < 2$, where four examples are specified to show the inpainting and denoising performance of the proposed method at 0.4, 0.8, 1.2 and 1.8, respectively.

vividly observe that the proposed method provides different performance with respect to p in terms of peak signal-to-noise (PSNR), where four results are obtained from $p = 0.4$, 0.8, 1.2, and 1.8, respectively.

IV. EXPERIMENTAL RESULTS

In this section, we evaluate the proposed ℓ_p-TTC method with extensive experiments on both simulated complex data and real visual data. The proposed method is benchmarked against APGL [40], FPCA [22], FP-LRTC [13], HaLRTC [8],

conventional PARAFAC [41] with full data and Cramér-Rao lower bound (CRLB) [42] for bistatic MIMO radar localization, and TCAM-TT [20], FaLRTC [8], SiLRTC [7], ℓ_p-PARAFAC [26] and IR-t-SVD [25] for color image inpainting and denoising, respectively.

A. Bistatic MIMO Radar Localization

In this section, the problem of source localization in bistatic MIMO radar is tackled via LRTC, where the direction of departure (DOD) and direction of arrival (DOA) are jointly estimated with the only available partial data at the front-end during multiple pulse periods. We consider the MIMO radar configuration with following parameters:

- Both transmit and receive arrays are uniform linear arrays (ULA) with M_t and M_r colocated antennas, respectively. d_t and d_r stand for the respective inter-element distances of antennas.
- K targets in the far-field, where $\{\theta_{mk}\}_{k=1}^K$ and $\{\phi_{nk}\}_{k=1}^K$ denote the DODs and DOAs with respect to the mth transmit and nth receive array normal, respectively;
- The transmit and receive steering matrices respectively are $\mathbf{A} = [\mathbf{a}(\theta_1), \cdots, \mathbf{a}(\theta_K)] \in \mathbb{C}^{M_t \times K}$ and $\mathbf{B} = [\mathbf{b}(\phi_1), \cdots, \mathbf{b}(\phi_K)] \in \mathbb{C}^{M_r \times K}$, with the steering vectors relative to ULA being $\mathbf{a}(\theta_k) = [1, e^{j2\pi d_t \sin(\theta_k)/\lambda}, \cdots, e^{j2\pi d_t(M_t-1)\sin(\theta_k)/\lambda}]^T$ and $\mathbf{b}(\phi_k) = [1, e^{j2\pi d_r \sin(\phi_k)/\lambda}, \cdots, e^{j2\pi d_r(M_r-1)\sin(\phi_k)/\lambda}]^T$.
- The coherent processing interval (CPI) consists of Q consecutive pulses and the radar cross section (RCS) is varying independently from pulse to pulse, where the RCS coefficients follow the Swerling II model [43]. $\mathbf{C} = [\mathbf{c}_1, \cdots, \mathbf{c}_Q]^T$ with $\mathbf{c}_q = [\gamma_{1q}, \cdots, \gamma_{Kq}]^T$ accounts for the Doppler effect and RCS fading;
- $\mathbf{S}_t = [\mathbf{s}_1, \cdots, \mathbf{s}_{M_t}]^T \in \mathbb{C}^{M_t \times L}$ holds the M_t narrowband pulse waveforms transmitted by the mth subarray, L being the number of samples per pulse period. The transmitted orthogonal waveforms are the columns of a Hadamard matrix.

Stacking Q pulses, $M_r \times M_t \times Q$ tensor $\mathcal{T}$ is obtained by following the PARAFAC model, details refer to [40], [44]:

$$\mathcal{T} = \mathcal{I}_{K \times 1}\mathbf{B}_{\times 2}\mathbf{A}_{\times 3}\mathbf{C} + \mathcal{N} \qquad (32)$$

$$\mathcal{X} = \mathcal{T}_\Omega \qquad (33)$$

where $\mathcal{I} \in \mathcal{C}^{K \times K \times K}$ is an identity tensor, $\mathcal{N}$ denotes the noise tensor, and $\mathcal{X}$ is the partially observed data. Given the restored $\mathcal{X}$, the matched-filter output can be obtained. Then the DODs and DOAs are computed via conventional parameter estimators, such as [45]. In our simulations, $M_t = M_r = 20$, $K = 5$, $Q = 256$, $L = 256$ and $\{\theta_k, \phi_k\}_{k=1}^K = \{\{10°, -30°\}, \{20°, -15°\}, \{30°, 5°\}, \{-10°, 15°\}, \{0°, 25°\}\}$. It is worth noting that since the simulated data for MIMO radar localization is complex, it should be transformed to real-valued data [46], and then restored by the proposed ℓ_p-TTC. Also, we can restore the target tensor from the real and imaginary parts, respectively. In addition, all algorithms are evaluated on degraded tensorial data with total missing ratio ρ at $\rho = 40\%$.
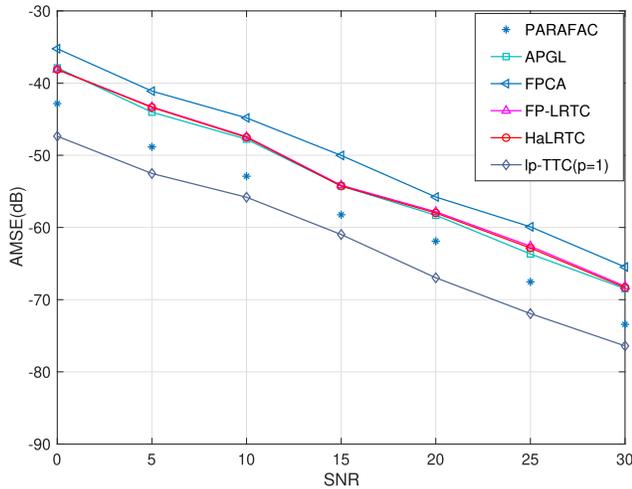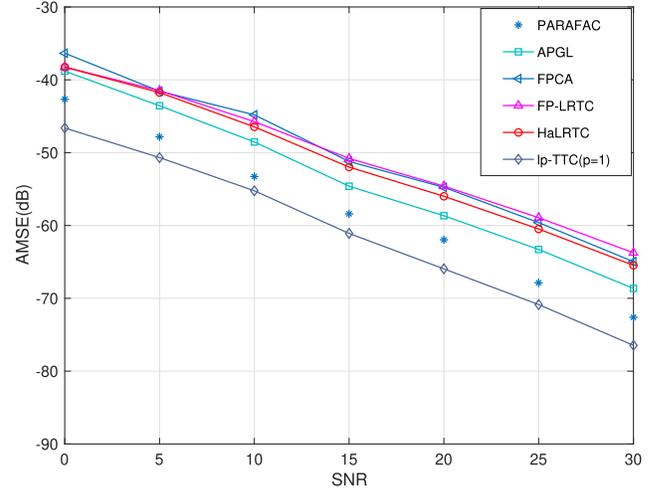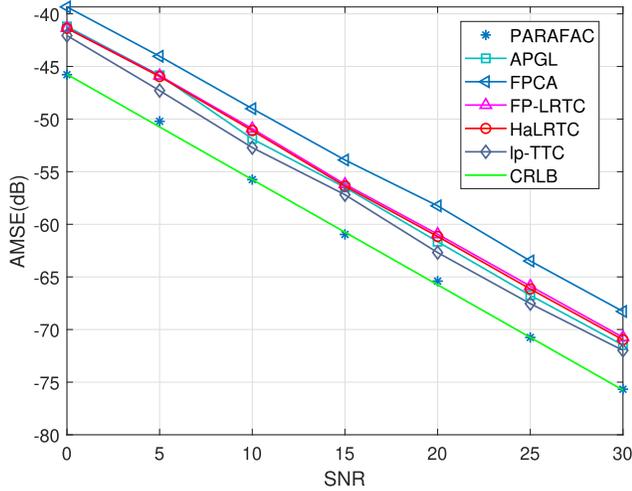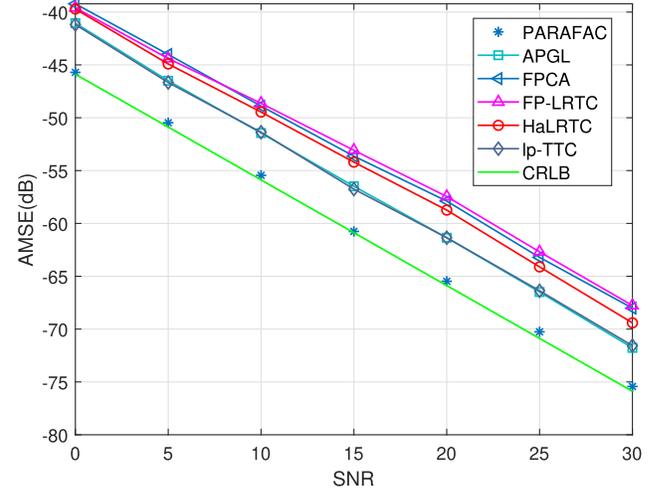
(a) AMSE versus SNR for DOD estimation　　　　(b) AMSE versus SNR for DOA estimation

Fig. 5. AMSE versus for different compared algorithms in Gaussian mixed model (GMM) noise, including APGL, FPGA, FP-LRTC, HaLRTC, and the proposed ℓ_p-TTC. Note that "Full data" means the result of conventional PARAFAC on the full data, and CRLB is not plotted here because it cannot be directly applied to the case of both missing entries and GMM noise.

(a) AMSE versus SNR for DOD estimation　　　　(b) AMSE versus SNR for DOA estimation

Fig. 6. AMSE versus SNR for different compared algorithms in Gaussian noise, where CRLB is compared as benchmark.

To measure the performance of the proposed method, the average mean square error (AMSE) between the estimated parameters and the original ones is used, which is defined as:

$$\text{AMSE} = \frac{1}{MK} \sum_{m=1}^{M} \sum_{k=1}^{K} (\xi_k - \hat{\xi}_k^m)^2 \qquad (34)$$

where M is the number of Monte Carlo trials, and $\hat{\xi}_k^m$ is the estimate of parameter ξ_k in the mth trial. Herein, ξ_k denotes the DOD θ_k or DOA ϕ_k, and $M = 100$.

In Fig. 5, the Swerling II model is chosen and we compare the proposed method versus state-of-the-art LRTC-based localization techniques via a Monte Carlo simulation, including APGL, FPGA, FP-LRTC, HaLRTC, the proposed ℓ_p-TTC at $p = 1$ and CRLB. Moreover, the performance of the conventional PARAFAC method [41] based on full data is provided as benchmark. In this experiment, the performance

of the proposed method robust against Gaussian mixture model (GMM) noise [47], [48] is investigated. GMM enables to well model the phenomenon in the presence of both Gaussian noise and outliers. Similar with [48], the ratio between variances of two-term Gaussian mixture noise is 100, signal-to-noise ratio (SNR) = 30 dB, and 10% noise samples are considered as outliers. From Fig. 5, we observe that the proposed method shows the best estimation performance among all approaches.

To further evaluate the performance of the proposed method, AMSE versus SNR is conducted in the presence of Gaussian noise [49], as shown in Fig. 6. The SNR is varied from 0 dB to 30 dB. Different with Fig. 5, the conventional PARAFAC method with full data is close to CRLB in Gaussian noise, which is also employed as the baseline in our work. For DOD estimation, it can be seen that the proposed ℓ_p-TTC outperforms all compared approaches, and achieves the best

| Original image | Missing and noisy image | lp-TTC | TTC | FaLRTC | SiLRTC | Ip-PARAFAC | IR-t-SVD |

Fig. 7. Visual comparison of different LRTC approaches on three color images: *Windows*, *Pillars* and *Peppers*, with 30% total missing and corrupted by 5dB salt-and-pepper noise. From left to right, the 1st column: ground truth; 2nd: incomplete and corrupted; 3rd: the proposed ℓ_p-TTC; 4th: TCAM-TT; 5th: FaLRTC; 6th: SiLRTC; 7th: ℓ_p-PARAFAC; Last: IR-t-SVD.

TABLE I

COMPARISON OF THE PROPOSED METHOD WITH STATE-OF-THE-ART TENSOR COMPLETION APPROACHES IN DIFFERENT MISSING RATIOS AND GMM NOISE, WHERE MSE IS APPLIED AS THE METRIC TO EVALUATE THE PERFORMANCE

Algorithms	30%	40%	50%	60%	70%
APGL	0.18151	0.23607	0.31104	0.33955	0.394
FPCA	0.17923	0.23411	0.282776	0.32965	0.384
FP-LRTC	0.32593	0.34415	0.36938	0.41883	0.46873
HaLRTC	0.33048	0.33856	0.36474	0.41429	0.46497
ℓ_p-TTC	**0.00975**	**0.007**	**0.00538**	**0.00453**	**0.0037**

performance among them for DOA estimation but is on par with APGL. Moreover, we observe that LRMC-based method, like FPCA, is inferior to LRTC-based ones.

In Table I, we compare the proposed method with numerous tensor completion approaches in different missing ratios from 30% to 70%, for bistatic MIMO radar localization. Unless otherwise specified, the parameter settings are the same as mentioned before. The quantitative results have clearly shown that the proposed method is superior to other compared methods, at least 2 order of magnitudes.

B. Color Image Inpainting and Denoising

Most of the existing image inpainting algorithms are developed for grayscale images. It is not trivial to extend them for color image inpainting since the noise statistics in R, G, and B channels can be very different for real noisy images except for the existence of outliers. In this section, simulation results are carried out to demonstrate the completion and denoising performance of the proposed ℓ_p-TTC

($p = 1$) on the standard color image set, including *Windows* ($300 \times 300 \times 3$), *Pillars* ($300 \times 300 \times 3$), *Peppers* ($317 \times 316 \times 3$), *House* ($225 \times 225 \times 3$), *Lenna* ($220 \times 220 \times 3$), *Building*2 ($300 \times 300 \times 3$), *Sea* ($300 \times 300 \times 3$), *Tree* ($300 \times 300 \times 3$), and *Shadow* ($360 \times 640 \times 3$).

Two objective measures, namely, PSNR and structural similarity (SSIM) [50] indexes, are adopted to provide quantitative and quality evaluations of the inpainting and denoising results. Since the PSNR is often inconsistent with human eye perception, even though it is the mostly used quality measure, the SSIM is employed to comprehensively reflect the performance. For the ground truth image $\mathcal{X}$ and the restored image $\hat{\mathcal{X}}$, the PSNR is defined as:

$$\text{PSNR} = 10\log_{10}\frac{255^2}{\text{MSE}} \qquad (35)$$

where 255 is the peak value of the color image under concern, with the mean squares error (MSE) given by

$$\text{MSE} = \frac{||\mathcal{X} - \hat{\mathcal{X}}||_F}{||\mathcal{X}||_F} \qquad (36)$$

Therefore, the smaller the MSE, the larger the PSNR, which implies the better performance.

The inpainting and denoising performance of the proposed method is examined with comparison to the state-of-the-art tensor completion approaches, including TTC [7], FaLRTC [8], SiLRTC [7], ℓ_p-PARAFAC [26] and IR-t-SVD [25]. The results are presented in Fig. 7, where 30% total entries are missing and the density of salt-and-pepper noise is 0.5. We can see that the proposed ℓ_p-TTC achieves the highest PSNR and SSIM on three color images, namely,

Fig. 8. The test images for Table II, namely, *House* (225 × 225 × 3), *Lenna* (220 × 220 × 3), *Building*2 (300 × 300 × 3), *Sea* (300 × 300 × 3), *Tree* (300 × 300 × 3), and *Shadow* (360 × 640 × 3).

TABLE II

COMPARISON OF THE PROPOSED METHOD WITH STATE-OF-THE-ART TENSOR COMPLETION APPROACHES IN DIFFERENT MISSING RATIOS WITH 2dB SALT-AND-PEPPER NOISE, WHERE MSE IS APPLIED AS THE METRIC TO EVALUATE THE PERFORMANCE

Images	Algorithms	30%	40%	50%	60%	70%
House	ℓ_p-TTC	**0.02599**	**0.01546**	**0.01315**	**0.01248**	**0.01208**
	TTC	0.0823	0.066	0.04946	0.04306	0.03629
	SiLRTC	0.09285	0.09921	0.10682	0.11825	0.13021
	FaLRTC	0.08563	0.09572	0.10573	0.11761	0.13074
	ℓ_p-PARAFAC	0.04755	0.0427	0.04156	0.03852	0.03125
Lenna	ℓ_p-TTC	**0.03629**	**0.03367**	**0.03251**	**0.03229**	**0.03007**
	TTC	0.08387	0.06406	0.05281	0.04679	0.04494
	SiLRTC	0.11711	0.12311	0.13064	0.14182	0.15743
	FaLRTC	0.10971	0.12012	0.12963	0.14132	0.15712
	ℓ_p-PARAFAC	0.09235	0.06011	0.05573	0.05533	0.05491
*Building*2	ℓ_p-TTC	**0.00316**	**0.003**	**0.00297**	**0.00293**	**0.00291**
	TTC	0.03323	0.02515	0.02031	0.01722	0.01582
	SiLRTC	0.05410	0.06493	0.07442	0.08534	0.09642
	FaLRTC	0.05161	0.06351	0.07322	0.08411	0.09532
	ℓ_p-PARAFAC	0.01042	0.00922	0.00912	0.00933	0.00912
Sea	ℓ_p-TTC	**0.00842**	**0.00478**	**0.00374**	**0.00349**	**0.0034**
	TTC	0.0528	0.04048	0.03523	0.02982	0.02604
	SiLRTC	0.06744	0.07455	0.08456	0.09646	0.10515
	FaLRTC	0.06361	0.07262	0.08425	0.09635	0.10464
	ℓ_p-PARAFAC	0.02053	0.01613	0.01571	0.01567	0.01539
Tree	ℓ_p-TTC	**0.02673**	**0.02239**	**0.01994**	**0.01913**	**0.01858**
	TTC	0.11046	0.0807	0.06625	0.05448	0.04813
	SiLRTC	0.13112	0.14101	0.15343	0.16823	0.18422
	FaLRTC	0.12623	0.13913	0.15233	0.16712	0.1833
	ℓ_p-PARAFAC	0.09683	0.06676	0.06631	0.0658	0.06563
Shadow	ℓ_p-TTC	**0.0103**	**0.00554**	**0.00467**	**0.0043**	**0.00418**
	TTC	0.07285	0.05603	0.04588	0.03864	0.03274
	SiLRTC	0.09042	0.10052	0.11662	0.13032	0.1478
	FaLRTC	0.08721	0.09933	0.11572	0.12963	0.14733
	ℓ_p-PARAFAC	0.03399	0.02372	0.02323	0.02294	0.02279

TABLE III

ROBUSTNESS OF THE PROPOSED METHOD CONDITIONED ON DIFFERENT LEVELS OF IMPULSIVE NOISE

	ℓ_p-TTC		TTC		FaLRTC		SiLRTC		ℓ_p-PARAFAC		IR-t-SVD	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
2dB	**25.5332**	**0.8156**	17.7496	0.2278	12.8591	0.0766	12.5244	0.0708	16.1935	0.2781	14.0141	0.1465
3dB	**25.9857**	**0.8337**	20.2460	0.4249	14.2347	0.1344	14.6430	0.1461	17.9288	0.4064	12.6154	0.1837
4dB	**26.1142**	**0.8381**	21.5423	0.5225	15.4014	0.1858	15.8552	0.2051	20.3557	0.5345	11.8688	0.1943
5dB	**26.1565**	**0.8540**	22.2978	0.5816	16.1879	0.2187	16.6566	0.2426	22.3567	0.5535	11.4637	0.2013
6dB	**26.2123**	**0.8586**	22.7432	0.6152	17.0601	0.2767	17.5461	0.3061	23.0152	0.5741	11.5198	0.2029

Windows, *Pillars* and *Peppers*. In addition, from the viewpoint of quantitative evaluation, a series of color images in Fig. 8 with different sizes are added into the simulation, and the results are tabulated in Table II. Consistent with the results in Table I, the proposed method also provides better performance in terms of MSE metric among all compared tensor completion approaches. We observe that TTC and ℓ_p-PARAFAC achieve comparable performance but still are inferior to ours. On the other hand, from Fig. 7 and Table II, it is verified that the proposed method can achieve

crystal clear image with strong structural integrity, and are showing good performance in terms of PSNR and MSE, respectively.

To further evaluate the robustness of the proposed method, we have conducted experiments on color image inpainting and denoising for different levels of impulsive noise, as tabulated in Table III. We observe that the proposed method achieves the best performance among all counterparts in terms of PSNR and SSIM metrics, which is consistent with our former analysis. The success of ℓ_p-TTC is attributed to the ℓ_p-regression
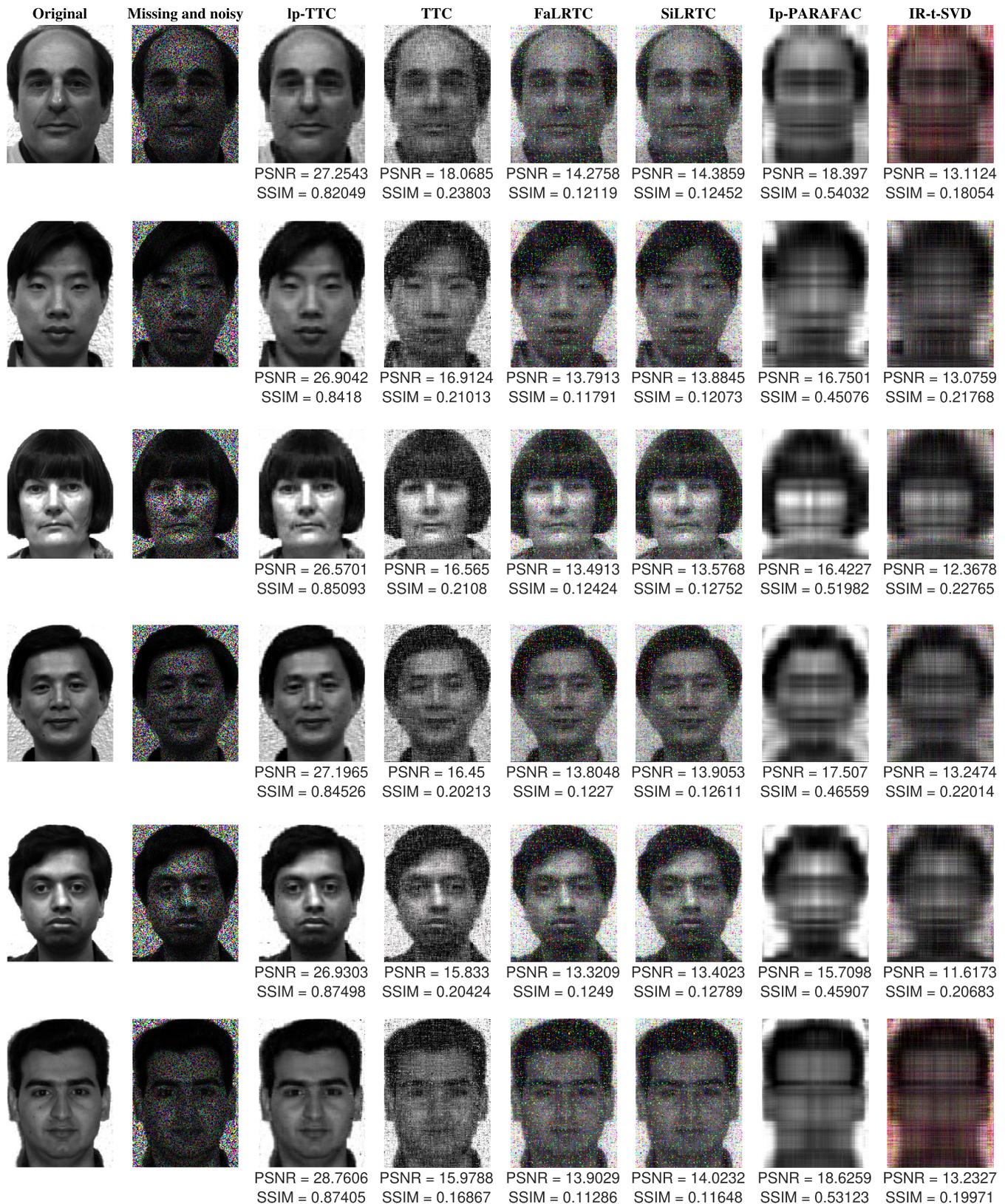
Fig. 9. Evaluation on Yale Face dataset by UCSD Computer Vision lab, where 50% missing entries and 6 dB salt and pepper noise are utilized to verify the robustness of the proposed method.

for non-Gaussian distribution. Moreover, in Figs. 9 and 10, we compare our method with several competitors on Yale Face dataset [51] and Flickr-Faces-HQ dataset [52], respectively, where 50% missing entries and 6 dB salt and pepper noise are added. Experimental results are presented, compared to other approaches, which vividly show the obtained image restored
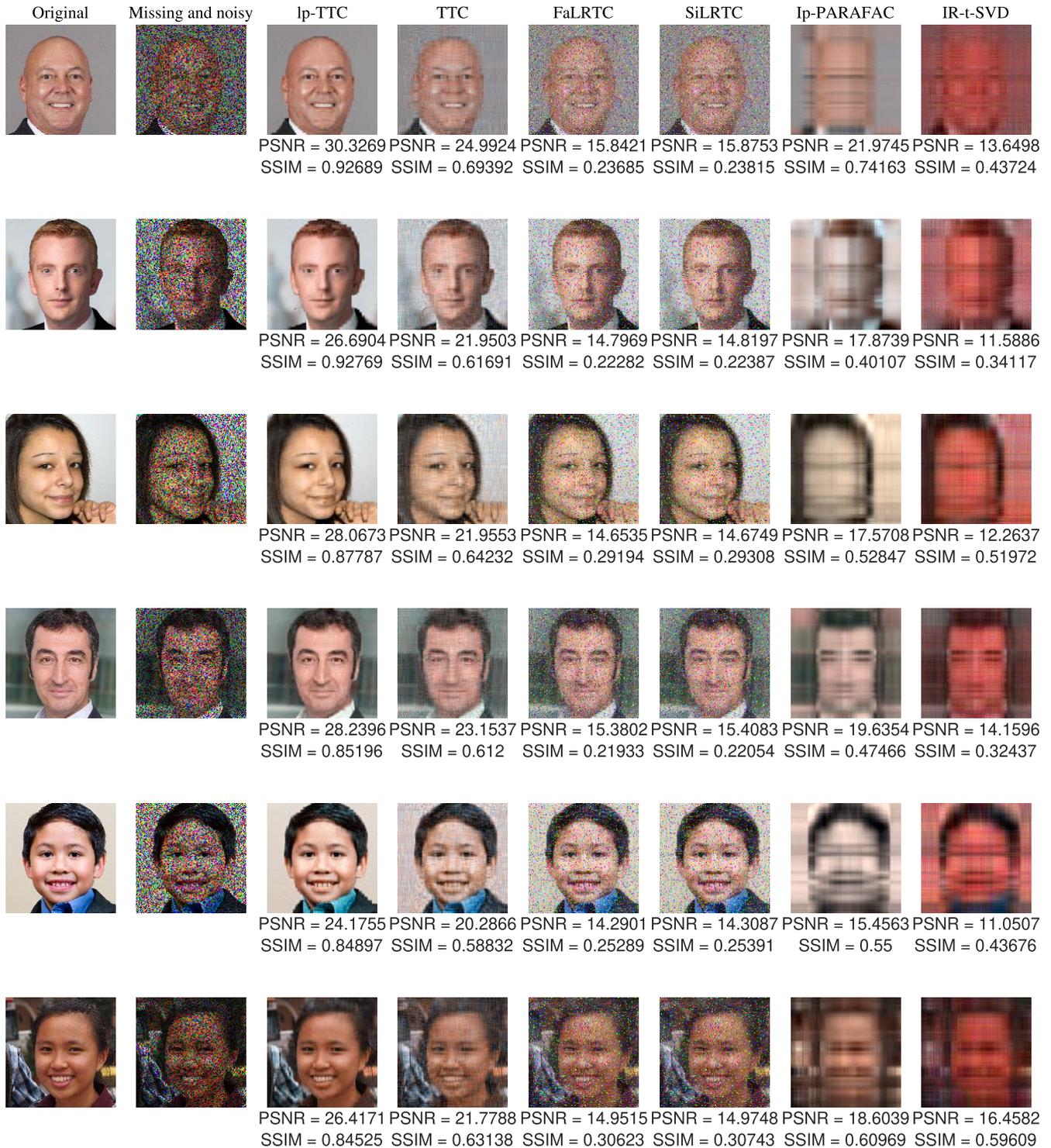
Fig. 10.   Evaluation results on Flicker-Faces-HQ dataset.

by the proposed method can achieve crystal clear image in both subjective and objective metrics.

V. CONCLUSION

In this work, a promising ℓ_p-TTC method is proposed to recover the missing entries and suppress the outliers from tensorial data. Motivated by the tensor factorization, a factored form of TT-format representation is applied into ℓ_p-regression minimization, which has the advantage of not requiring to pre-define the best user-defined weights in TT rank model. The proposed method succeeds to restore the target tensor from both missing entries and outliers corrupted observations. The effectiveness and superiority of the proposed scheme is verified by the applications of bistatic MIMO radar localization and

color image inpainting and denoising tests, from simulated to visual data, respectively. However, although the proposed method is empirically convergent, the theoretical analysis of convergence is still an open problem.

REFERENCES

[1] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.

[2] Y. Liu, Z. Long, H. Huang, and C. Zhu, "Low CP rank and Tucker rank tensor completion for estimating missing components in image data," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 4, pp. 944–954, Apr. 2020.

[3] M. Steinlechner, "Riemannian optimization for high-dimensional tensor completion," *SIAM J. Sci. Comput.*, vol. 38, no. 5, pp. S461–S484, Jan. 2016.

[4] S. Javed, A. Mahmood, T. Bouwmans, and S. K. Jung, "Spatiotemporal low-rank modeling for complex scene background initialization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 6, pp. 1315–1329, Jun. 2018.

[5] A. Cichocki *et al.*, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 145–163, Mar. 2015.

[6] F. Cao, M. Cai, and Y. Tan, "Image interpolation via low-rank matrix completion and recovery," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 8, pp. 1261–1270, Aug. 2015.

[7] J. A. Bengua, H. N. Phien, H. D. Tuan, and M. N. Do, "Efficient tensor completion for color image and video recovery: Low-rank tensor train," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2466–2479, May 2017.

[8] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, Jan. 2013.

[9] T. Franz, A. Schultz, S. Sizov, and S. Staab, "TripleRank: Ranking semantic web data by tensor decomposition," in *The Semantic Web—ISWC*, A. Bernstein *et al.*, Eds. Berlin, Germany: Springer, 2009, pp. 213–228.

[10] H. Tan, B. Cheng, W. Wang, Y.-J. Zhang, and B. Ran, "Tensor completion via a multi-linear low-n-rank factorization model," *Neurocomputing*, vol. 133, pp. 161–169, Jun. 2014.

[11] H. Becker, L. Albera, P. Comon, R. Gribonval, F. Wendling, and I. Merlet, "Brain-source imaging: From sparse to tensor models," *IEEE Signal Process. Mag.*, vol. 32, no. 6, pp. 100–112, Nov. 2015.

[12] Q. Zhao, L. Zhang, and A. Cichocki, "Bayesian CP factorization of incomplete tensors with automatic rank determination," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1751–1763, Sep. 2015.

[13] L. Yang, Z.-H. Huang, and X. Shi, "A fixed point iterative method for low n-rank tensor pursuit," *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2952–2962, Jun. 2013.

[14] Z. Zhang, G. Ely, S. Aeron, N. Hao, and M. Kilmer, "Novel methods for multilinear data completion and de-noising based on tensor-SVD," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Columbus, OH, USA, Jun. 2014, pp. 3842–3849.

[15] T. Sadowski and R. Zdunek, "Image completion with filtered alternating least squares Tucker decomposition," in *Proc. Signal Processing: Algorithms, Archit., Arrangements, Appl. (SPA)*, Sep. 2019, pp. 241–245.

[16] J. Xue, Y. Zhao, W. Liao, J. C. Chan, and S. G. Kong, "Enhanced sparsity prior model for low-rank tensor completion," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 11, pp. 4567–4581, Nov. 2020.

[17] L. Yang, J. Fang, H. Li, and B. Zeng, "An iterative reweighted method for Tucker decomposition of incomplete tensors," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4817–4829, Sep. 2016.

[18] Y. Xu, Z. Wu, J. Chanussot, and Z. Wei, "Hyperspectral computational imaging via collaborative Tucker3 tensor decomposition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 1, pp. 98–111, Jan. 2021.

[19] I. V. Oseledets, "Tensor-train decomposition," *SIAM J. Sci. Comput.*, vol. 33, no. 5, pp. 2295–2317, Jan. 2011.

[20] W. Wang, V. Aggarwal, and S. Aeron, "Tensor completion by alternating minimization under the tensor train (TT) model," 2016, *arXiv:1609.05587*. [Online]. Available: http://arxiv.org/abs/1609.05587

[21] Y.-L. Chen, C.-T. Hsu, and H.-Y. M. Liao, "Simultaneous tensor decomposition and completion using factor priors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 3, pp. 577–591, Mar. 2014.

[22] S. Ma, D. Goldfarb, and L. Chen, "Fixed point and Bregman iterative methods for matrix rank minimization," *Math. Program.*, vol. 128, nos. 1–2, pp. 321–353, 2011.

[23] I. Kajo, N. Kamel, and Y. Ruichek, "Incremental tensor-based completion method for detection of stationary foreground objects," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 5, pp. 1325–1338, May 2019.

[24] M. R. Khokher, A. Bouzerdoum, and S. L. Phung, "A super descriptor tensor decomposition for dynamic scene recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 4, pp. 1063–1076, Apr. 2019.

[25] W. Sun, X. Lin, H. C. So, L. Huang, and Q. Li, "Iteratively reweighted tensor SVD for robust multi-dimensional harmonic retrieval," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Shanghai, China, Mar. 2016, pp. 4318–4322.

[26] X. Lin, L. Huang, and W. Sun, "ℓ_p-PARAFAC for joint DOD and DOA estimation in bistatic MIMO radar," in *Proc. CIE Int. Conf. Radar (RADAR)*, Guangzhou, China, 2016, pp. 1–5.

[27] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge, U.K.: Cambridge Univ. Press, Jun. 2010.

[28] S. Holtz, T. Rohwedder, and R. Schneider, "On manifolds of tensors of fixed TT-rank," *Numer. Math.*, vol. 120, no. 4, pp. 701–731, Apr. 2012.

[29] A. Cichocki, N. Lee, I. Oseledets, A.-H. Phan, Q. Zhao, and D. P. Mandic, "Tensor networks for dimensionality reduction and large-scale optimization: Part 1 low-rank tensor decompositions," *Found. Trends Mach. Learn.*, vol. 9, nos. 4–5, pp. 249–429, 2016.

[30] A. Cichocki *et al.*, "Tensor networks for dimensionality reduction and large-scale optimization: Part 2 applications and future perspectives," *Found. Trends Mach. Learn.*, vol. 9, no. 6, pp. 249–429, 2017.

[31] A. Karami, M. Yazdi, and A. Z. Asli, "Noise reduction of hyperspectral images using kernel non-negative Tucker decomposition," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 487–493, Jun. 2011.

[32] B. Jiang, S. Ma, and S. Zhang, "Low-M-rank tensor completion and robust tensor PCA," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 6, pp. 1390–1404, Dec. 2018.

[33] J. Yang, Y. Zhu, K. Li, J. Yang, and C. Hou, "Tensor completion from structurally-missing entries by low-TT-rankness and fiber-wise sparsity," *IEEE J. Sel. Topics Signal Process*, vol. 12, no. 6, pp. 1420–1434, Dec. 2018.

[34] G. Vidal, "Efficient simulation of one-dimensional quantum many-body systems," *Phys. Rev. Lett.*, vol. 93, no. 4, Jul. 2004, Art. no. 040502.

[35] P. D. G. Mary, B. Madathil, and S. N. George, "Entropy-based reweighted tensor completion technique for video recovery," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 2, pp. 415–426, Feb. 2020.

[36] Y. Wang, W. Yin, and J. Zeng, "Global convergence of ADMM in nonconvex nonsmooth optimization," *J. Sci. Comput.*, vol. 78, no. 1, pp. 29–63, Jun. 2018.

[37] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2479–2493, Jul. 2009.

[38] G. Marjanovic and V. Solo, "On ℓ_q optimization and matrix completion," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5714–5724, Nov. 2012.

[39] W.-J. Zeng and H. C. So, "Outlier-robust matrix completion via ℓ_p-minimization," *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1125–1140, Dec. 2017.

[40] L.-T. Huang, A. L. F. de Almeida, and H. C. So, "Target estimation in bistatic MIMO radar via tensor completion," *Signal Process.*, vol. 120, pp. 654–659, Mar. 2016.

[41] R. Bro, "PARAFAC. Tutorial and applications," *Chemometrics Intell. Lab. Syst.*, vol. 38, no. 2, pp. 149–171, Oct. 1997.

[42] M. La Manna and D. R. Fuhrmann, "Cramér–Rao lower bounds comparison for 2D hybrid-MIMO and MIMO radar," *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 2, pp. 404–413, Mar. 2017.

[43] T. Aittomaki and V. Koivunen, "Performance of MIMO radar with angular diversity under swerling scattering models," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 1, pp. 101–114, Feb. 2010.

[44] S. Q. Sun, W. U. Bajwa, and A. P. Petropulu, "MIMO-MC radar: A MIMO radar approach based on matrix completion," *IEEE Trans. Aerosp. Electron. Sys.*, vol. 51, no. 3, pp. 1839–1852, Jul. 2015.

[45] X. Wang, W. Wang, J. Liu, Q. Liu, and B. Wang, "Tensor-based real-valued subspace approach for angle estimation in bistatic MIMO radar with unknown mutual coupling," *Signal Process.*, vol. 116, pp. 152–158, Nov. 2015.

[46] Q. Liu, W. Wang, D. Liang, and X. Wang, "Real-valued reweighted l_1 norm minimization method based on data reconstruction in MIMO radar," *IEICE Trans. Commun.*, vol. E98.B, no. 11, pp. 2307–2313, 2015.

[47] M. Rangaswamy, D. Weiner, and A. Ozturk, "Computer generation of correlated non-Gaussian radar clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 1, pp. 106–116, Jan. 1995.

[48] Q. Liu, Y. Gu, and H. C. So, "DOA estimation in impulsive noise via low-rank matrix approximation and weakly convex optimization," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 6, pp. 3603–3616, Dec. 2019.

[49] J. Liang, D. Wang, L. Su, B. Chen, H. Chen, and H. C. So, "Robust MIMO radar target localization via nonconvex optimization," *Signal Process.*, vol. 122, pp. 33–38, May 2016.

[50] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[51] A. S. Georghiades, P. N. Belhumeur, and D. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 643–660, Jun. 2001.

[52] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 4396–4405.

Xiaopeng Li (Member, IEEE) was born in Hebei, China. He received the B.E. degree in electronic science and technology from Yanshan University, Qinhuangdao, China, in 2015, and the M.Sc. degree in electronic information engineering from the City University of Hong Kong, China, in 2018, where he is currently pursuing the Ph.D. degree with the Department of Electrical Engineering. From 2018 to 2019, he was a Research Assistant with the College of Information Engineering, Shenzhen University, China. His research interests include robust signal processing and sparse approximation.

Hui Cao (Member, IEEE) received the Ph.D. degree in electronic engineering from the City University of Hong Kong in October 2017. He is currently a Lecturer with the School of Information Engineering, Wuhan University of Technology. His research interests include statistical signal processing, parameter estimation, and source localization.

Qi Liu (Member, IEEE) received the B.E. degree in measuring and control technology and instrumentations, and the M.E. degree in control science and engineering from Harbin Engineering University, Harbin, China, in 2013 and 2016, respectively, and the Ph.D. degree in electrical engineering from the City University of Hong Kong, Hong Kong, China, in 2019.

From 2018 to 2019, he was a Visiting Scholar at the Department of Electrical and Computer Engineering, University of California at Davis, Davis, CA, USA. During this period, he was invited to give talks at the University of California at San Diego, San Diego, and California Institute of Technology. Currently, he works as a Research Fellow with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore. His research interests broadly lie in multidimensional signal processing, optimization methods, and neuromorphic computing, with applications to DOA estimation, robust sparse recovery, spiking neural networks, matrix completion, keyword spotting, and image denoising and inpainting, as well as MIMO radar. He was a recipient of the Best Paper Award at the 2019 IEEE International Conference on Signal, Information and Data Processing. He serves as a reviewer for more than 20 major referred journals and conference papers.

Yuntao Wu (Senior Member, IEEE) received the Ph.D. degree in information and communication engineering from the National Key Laboratory for Radar Signal Processing, Xidian University, in December 2003. From April 2004 to September 2006, he was a Post-Doctoral Researcher at the Institute of Acoustics, Chinese Academy of Sciences, Beijing, China. During October 2006 and February 2008, he worked as a Senior Research Fellow with the City University of Hong Kong. From April 2013 to March 2014, he was a Visiting Researcher with the Faculty of Engineering, Bar-Ilan University, Israel. Currently, he is a full-time Professor with Wuhan Institute of Technology. At the same time, he is also working on a Chutian Scholar Project in Hubei as a Distinguish Professor at the same university. He has published over 80 journals and conference papers. His research interests include signal detection, parameter estimation in array signal processing and source localization for wireless sensor networks, and biomedicine signal analysis. He is a Senior Member of the Chinese Institute of Electronic Engineers (CIE) and an Associate Editor of *Multidimensional Systems and Signal Processing* (Springer).