ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-7281-7605-5/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICASSP39728.2021.9415104

RANK-REVEALING BLOCK-TERM DECOMPOSITION FOR TENSOR COMPLETION

Athanasios A. Rontogiannis[1] *Paris V. Giampouras*[2] *Eleftherios Kofidis*[3]

[1]IAASARS, National Observatory of Athens, 152 36 Penteli, Greece
tronto@noa.gr
[2]Mathematical Inst. for Data Science, Johns Hopkins University, Baltimore, MD 21218, USA
parisg@jhu.edu
[3]Dept. of Statistics and Insurance Science, University of Piraeus, 185 34 Piraeus, Greece
kofidis@unipi.gr

ABSTRACT

The so-called block-term decomposition (BTD) tensor model has been recently receiving increasing attention due to its enhanced ability of representing systems and signals that are composed of *blocks* of rank higher than one, a scenario encountered in numerous and diverse applications. In this paper, BTD is employed for the completion of a tensor from its partially observed entries. A novel method is proposed, which is based on the idea of imposing column sparsity *jointly* on the BTD factors and in a *hierarchical* manner. This way the number of block terms and their ranks can also be estimated, as the numbers of factor columns of non-negligible magnitude. Following a block successive upper bound minimization (BSUM) approach with appropriate choice of the surrogate majorizing functions is shown to result in an alternating hierarchical iteratively reweighted least squares (HIRLS) algorithm, which is fast converging and enjoys high computational efficiency, as it relies in its iterations on small-sized sub-problems with closed-form solutions. Simulation results with both synthetic and real data are reported, which demonstrate the effectiveness of the proposed scheme.

Index Terms— Block successive upper bound minimization (BSUM), block-term tensor decomposition (BTD), completion, rank, tensor

1. INTRODUCTION

The ubiquity in our big data era of data with (explicitly or implicitly) multiple dimensions/relations that are often incomplete and/or uncertain has given rise to numerous applications of tensor completion [1], such as image and video in-painting [2], hyperspectral imaging [3], prediction of multi-dimensional non-stationary wireless channels [4], semiconductor manufacturing [5], and computational materials science [6], to name only a few. Though still lacking in sufficient theoretical foundations and algorithmic variety when compared to its matrix-based counterpart, tensor completion has already a quite rich literature [1], which includes (among other approaches) methods based on low-rank decomposition models. In addition to the classical Canonical Polyadic Decomposition (CPD) and/or Tucker decomposition [7, 8, 9, 10], less well-known models, such as t-SVD [11], tensor trains [12] (and tensor rings [2]), and M-decomposition [13], have been studied in the context of the completion problem. The model order (e.g., the tensor rank in CPD) is almost always assumed *a-priori* known. Exceptions include [7, 8] (see [1] for additional references) where the CPD rank is also estimated in the course of the completion through an appropriate regularization of the tensor reconstruction cost function.

The so-called block-term decomposition (BTD) [14] has been recently receiving increasing attention due to its enhanced ability of representing systems and signals that are composed of *blocks* of rank higher than one, a scenario encountered in numerous and diverse applications (see, e.g., [15] and references therein). The most popular BTD model, namely the rank-$(L_r, L_r, 1)$ decomposition, is defined as follows for an $I \times J \times K$ tensor $\boldsymbol{\mathcal{X}}$:

$$\boldsymbol{\mathcal{X}} = \sum_{r=1}^{R} \mathbf{E}_r \circ \mathbf{c}_r, \tag{1}$$

where $\mathbf{E}_r$ is an $I \times J$ matrix of rank L_r, $\mathbf{c}_r$ is a nonzero column K-vector and $\circ$ denotes outer product. Clearly, $\mathbf{E}_r$ can be written as a matrix product $\mathbf{A}_r\mathbf{B}_r^{\mathrm{T}}$ with the matrices $\mathbf{A}_r = \begin{bmatrix} \mathbf{a}_{r1} & \mathbf{a}_{r2} & \cdots & \mathbf{a}_{rL_r} \end{bmatrix} \in \mathbb{C}^{I\times L_r}$ and $\mathbf{B}_r = \begin{bmatrix} \mathbf{b}_{r1} & \mathbf{b}_{r2} & \cdots & \mathbf{b}_{rL_r} \end{bmatrix} \in \mathbb{C}^{J\times L_r}$ being of full column rank, L_r. To the best of our knowledge, BTD-based completion has been so far only studied in [16], in the context of spectrum cartography, where joint decomposition and completion accomplish the so-called dis-aggregation task. For the case of the observed tensor entries being randomly sam-

P. V. Giampouras is supported by the European Union under the Horizon 2020 Marie Skłodowska-Curie Global Fellowship program: HyPPOCRATES — H2020-MSCA-IF-2018, Grant Agreement Number: 844290.

The work of E. Kofidis has been partly supported by the University of Piraeus Research Center.

pled in a uniform manner, of interest also in this paper, an $I \times J \times K$ tensor $\boldsymbol{\mathcal{Y}}$ is completed from its partially observed version $\boldsymbol{\mathcal{Z}} \triangleq \boldsymbol{\mathcal{W}} * \boldsymbol{\mathcal{Y}}$, where $\boldsymbol{\mathcal{W}}$ is a binary sampling tensor with ones at the observed positions and zeros elsewhere and $*$ is the Hadamard product, by solving

$$\min_{\mathbf{A},\mathbf{B},\mathbf{C}} f(\mathbf{A},\mathbf{B},\mathbf{C}) \triangleq \frac{1}{2}\left\| \boldsymbol{\mathcal{Z}} - \boldsymbol{\mathcal{W}} * \left(\sum_{r=1}^{R} \mathbf{A}_r \mathbf{B}_r^{\mathrm{T}} \circ \mathbf{c}_r \right) \right\|_{\mathrm{F}}^2 \tag{2}$$

with respect to (w.r.t.) $\mathbf{A} \triangleq \left[\begin{array}{cccc} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_R \end{array} \right]$, $\mathbf{B} \triangleq \left[\begin{array}{cccc} \mathbf{B}_1 & \mathbf{B}_2 & \cdots & \mathbf{B}_R \end{array} \right]$, and $\mathbf{C} \triangleq \left[\begin{array}{cccc} \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_R \end{array} \right]$. In [16], and in order to avoid overfitting, the following regularized version of the problem is considered,

$$\min_{\mathbf{A},\mathbf{B},\mathbf{C}} f(\mathbf{A},\mathbf{B},\mathbf{C}) + \lambda_1 \|\mathbf{A}\|_{\mathrm{F}}^2 + \lambda_2 \|\mathbf{B}\|_{\mathrm{F}}^2 + \lambda_3 \|\mathbf{C}\|_{\mathrm{F}}^2, \tag{3}$$

with λ_i, $i = 1, 2, 3$ being scaling parameters (finally chosen to be all equal). Exact block coordinate descent (BCD) is adopted for its solution. The number of block terms, R, and their ranks, L_r, are assumed known in the algorithm, an assumption that can be unrealistic in practice and/or require computationally intensive trial-and-error search.

Recently, joint BTD model selection and computation for complete tensors was given special attention (see [15] and references therein). A new method was developed and reported in [17, 15], based on the idea of imposing column sparsity *jointly* on the factors in a *hierarchical* manner that matches the model structure, using $\ell_{1,2}$ norm regularization. This allows to also estimate the number of block terms (R) and their ranks (L_r) as the numbers of factor columns of non-negligible magnitude [17, 15]. Promising results were obtained with both synthetic and real data. Analogous ideas are adopted here to devise a BTD-based completion scheme. Following a block successive upper bound minimization (BSUM) approach with appropriate choice of the surrogate majorizing functions is shown to result in an alternating hierarchical iteratively reweighted least squares (HIRLS) algorithm, which is fast converging and enjoys high computational efficiency, as it relies in its iterations on small-sized sub-problems with closed-form solutions. Simulation results with both synthetic and real data are reported, which demonstrate the effectiveness of the proposed scheme.

2. PROBLEM STATEMENT

Our aim is to complete a (noisy) $I \times J \times K$ randomly sampled tensor $\boldsymbol{\mathcal{Y}}$ through its best BTD approximation as in (2), where the ranks R and L_r, $r = 1, 2, \ldots, R$ are considered *a-priori* unknown. As suggested by our earlier work [17, 15], model selection can be effected by adding to the cost of (2) an $\ell_{1,2}$-type regularization term that penalizes high ranks and perfectly matches the structure of the BTD model.[1] Thus, we propose to solve the following modification of (2)

$$\min_{\mathbf{A},\mathbf{B},\mathbf{C}} f(\mathbf{A},\mathbf{B},\mathbf{C}) + \lambda \|\mathbf{F}(\mathbf{A},\mathbf{B},\mathbf{C})\|_{1,2}, \tag{4}$$

where λ is a regularization parameter and the $2 \times R$ matrix $\mathbf{F}(\mathbf{A},\mathbf{B},\mathbf{C})$ is defined as

$$\mathbf{F}(\mathbf{A},\mathbf{B},\mathbf{C}) \triangleq \left[\begin{array}{cccc} \|\mathbf{G}_1\|_{1,2} & \|\mathbf{G}_2\|_{1,2} & \cdots & \|\mathbf{G}_R\|_{1,2} \\ \|\mathbf{c}_1\|_2 & \|\mathbf{c}_2\|_2 & \cdots & \|\mathbf{c}_R\|_2 \end{array} \right]. \tag{5}$$

$\mathbf{G} \triangleq \left[\begin{array}{cc} \mathbf{A}^{\mathrm{T}} & \mathbf{B}^{\mathrm{T}} \end{array} \right]^{\mathrm{T}}$ is the $(I + J) \times \sum_{r=1}^{R} L_r$ matrix resulting from stacking the factors $\mathbf{A}$ and $\mathbf{B}$ and $\mathbf{G}_r \triangleq \left[\begin{array}{cc} \mathbf{A}_r^{\mathrm{T}} & \mathbf{B}_r^{\mathrm{T}} \end{array} \right]^{\mathrm{T}}$ denote its rth $(I + J) \times L_r$ block. From (4) and (5) we see that the proposed penalty term is composed of two $\ell_{1,2}$ norms combined hierarchically. The column sparsity promoting property of the $\ell_{1,2}$ norm [18] is taken advantage of at two levels. At the upper level, the minimization of the $\ell_{1,2}$ norm of $\mathbf{F}(\mathbf{A},\mathbf{B},\mathbf{C})$ favors the elimination of columns of $\mathbf{C}$ and corresponding whole blocks of $\mathbf{A}$ and $\mathbf{B}$ (via the minimization of the norms of the $\mathbf{c}_r$'s and those of the corresponding $\mathbf{G}_r$'s), while, at the lower level, column sparsity (hence low rank) is imposed on the surviving $\mathbf{A}_r$'s and $\mathbf{B}_r$'s jointly, through the $\ell_{1,2}$ norms of the $\mathbf{G}_r$'s. Thus, by overestimating the unknown ranks as R_{ini} and L_{ini} and properly selecting the regularization parameter λ, the actual tensor structure can be found by the previously described elimination procedure. An efficient algorithm for (4) that implements the above idea is presented in the next section.

It will be useful in our development to recall the factorization of the mode unfoldings of a BTD tensor $\boldsymbol{\mathcal{X}} \triangleq \sum_{r=1}^{R} \mathbf{A}_r \mathbf{B}_r^{\mathrm{T}} \circ \mathbf{c}_r$ [14]:

$$\mathbf{X}_{(1)}^{\mathrm{T}} = (\mathbf{B} \odot \mathbf{C})\mathbf{A}^{\mathrm{T}} \triangleq \mathbf{P}\mathbf{A}^{\mathrm{T}}, \tag{6}$$

$$\mathbf{X}_{(2)}^{\mathrm{T}} = (\mathbf{C} \odot \mathbf{A})\mathbf{B}^{\mathrm{T}} \triangleq \mathbf{Q}\mathbf{B}^{\mathrm{T}}, \tag{7}$$

$$\begin{aligned} \mathbf{X}_{(3)}^{\mathrm{T}} &= \left[\begin{array}{c} (\mathbf{A}_1 \odot_{\mathrm{c}} \mathbf{B}_1)\mathbf{1}_{L_1} \cdots (\mathbf{A}_R \odot_{\mathrm{c}} \mathbf{B}_R)\mathbf{1}_{L_R} \end{array} \right] \mathbf{C}^{\mathrm{T}} \\ &\triangleq \mathbf{S}\mathbf{C}^{\mathrm{T}}, \end{aligned} \tag{8}$$

where $\odot$ stands for the Khatri-Rao product in its general (partition-wise) version and $\odot_{\mathrm{c}}$ is its column-wise version. The all ones column N-vector is denoted by $\mathbf{1}_N$.

3. PROPOSED METHOD

First, we rewrite (4) explicitly w.r.t. the BTD factors, and add a small positive constant η^2 to ensure the smoothness of the objective function:

$$\begin{aligned} \min_{\mathbf{A},\mathbf{B},\mathbf{C}} \frac{1}{2} \left\| \boldsymbol{\mathcal{Z}} - \boldsymbol{\mathcal{W}} * \left(\sum_{r=1}^{R} \mathbf{A}_r \mathbf{B}_r^{\mathrm{T}} \circ \mathbf{c}_r \right) \right\|_{\mathrm{F}}^2 + \\ \lambda \sum_{r=1}^{R} \sqrt{ \left(\sum_{l=1}^{L} \sqrt{ \|\mathbf{a}_{rl}\|_2^2 + \|\mathbf{b}_{rl}\|_2^2 + \eta^2 } \right)^2 + \|\mathbf{c}_r\|_2^2 + \eta^2 }, \end{aligned} \tag{9}$$

[1]The mixed $\ell_{1,2}$ norm of a matrix is defined as the ℓ_1 norm of the vector of the ℓ_2 norms of its columns [18].

where R and L are assumed to be (probably loose) upper bounds of the ranks. For the non-convex problem (9) we propose an inexact BCD procedure[2], employing a majorization-minimization solver at each block minimization step. It falls in the block successive upper bound minimization (BSUM) framework [20]. In our case, the block variables are the BTD factors $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, for which appropriate surrogate functions (of the quadratic type [20]) will be defined and successively minimized.

The objective function of the $\mathbf{A}$ sub-problem of (9) can be expressed at the kth BCD iteration as

$$f_{\mathbf{A}}(\mathbf{A}|\mathbf{B}^k,\mathbf{C}^k) = \frac{1}{2}\left\|\mathbf{Z}_{(1)}^{\mathrm{T}} - \mathbf{W}_{(1)}^{\mathrm{T}} * (\mathbf{P}^k\mathbf{A}^{\mathrm{T}})\right\|_{\mathrm{F}}^2 + \lambda\sum_{r=1}^{R}\sqrt{\left(\sum_{l=1}^{L}\sqrt{\|\mathbf{a}_{rl}\|_2^2 + \|\mathbf{b}_{rl}^k\|_2^2+\eta^2}\right)^2 + \|\mathbf{c}_r^k\|_2^2+\eta^2}, \tag{10}$$

where the tensors $\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Z}}$ and $\boldsymbol{\mathcal{W}}$ are seen in their mode-1 unfoldings. To allow this sub-problem to have closed-form solution for $\mathbf{A}$, we define a local tight upper bound function of (10) as a rough second-order Taylor approximation of $f_{\mathbf{A}}(\mathbf{A}|\mathbf{B}^k,\mathbf{C}^k)$ around $\mathbf{A}^k$:

$$g_{\mathbf{A}}(\mathbf{A}|\mathbf{A}^k,\mathbf{B}^k,\mathbf{C}^k) = f_{\mathbf{A}}(\mathbf{A}^k|\mathbf{B}^k,\mathbf{C}^k) + \mathrm{tr}\{(\mathbf{A}-\mathbf{A}^k) \nabla_{\mathbf{A}} f_{\mathbf{A}}(\mathbf{A}^k|\mathbf{B}^k,\mathbf{C}^k)\} + \tfrac{1}{2}\mathrm{vec}(\mathbf{A}-\mathbf{A}^k)^{\mathrm{T}}\bar{\mathbf{H}}_{\mathbf{A}^k}\mathrm{vec}(\mathbf{A}-\mathbf{A}^k),$$

where $\mathrm{vec}(\cdot)$ is the row vectorization operator and the $ILR \times ILR$ approximate Hessian matrix $\bar{\mathbf{H}}_{\mathbf{A}^k}$ is given by

$$\bar{\mathbf{H}}_{\mathbf{A}^k} \triangleq \mathbf{I}_I \otimes (\mathbf{P}^{k\mathrm{T}}\mathbf{P}^k + \lambda\mathbf{D}^k). \tag{11}$$

As explained below, the choice of an appropriate approximate Hessian matrix, as in (11), is of crucial importance for the convergence and complexity of the algorithm. In (11), $\mathbf{D}^k \triangleq (\mathbf{D}_1^k \otimes \mathbf{I}_L)\mathbf{D}_2^k$ is a diagonal *reweighting* matrix composed of the $R\times R$ diagonal matrix $\mathbf{D}_1^k$ and the $RL\times RL$ diagonal matrix $\mathbf{D}_2^k$. The rth diagonal entry of $\mathbf{D}_1^k$ is

$$\mathbf{D}_1^k(r,r) = \left[\left(\sum_{l=1}^{L}\sqrt{\|\mathbf{a}_{rl}^k\|_2^2 + \|\mathbf{b}_{rl}^k\|_2^2+\eta^2}\right)^2 + \|\mathbf{c}_r^k\|_2^2+\eta^2\right]^{-1/2} \tag{12}$$

and the $((r-1)L+l)$th diagonal entry of $\mathbf{D}_2^k$ is given by

$$\mathbf{D}_2^k((r-1)L+l,(r-1)L+l) = (\|\mathbf{a}_{rl}^k\|_2^2 + \|\mathbf{b}_{rl}^k\|_2^2+\eta^2)^{-1/2}. \tag{13}$$

Note that $\mathbf{D}_1^k$ and $\mathbf{D}_2^k$ refer to the first and second elimination steps, respectively, described in the previous section. Minimizing $g_{\mathbf{A}}$ yields the following *closed-form* solution:

$$\mathbf{A}^{k+1} = \mathbf{A}^k + [(\mathbf{Z}_{(1)} - \mathbf{W}_{(1)} * (\mathbf{A}^k\mathbf{P}^{k\mathrm{T}}))\mathbf{P}^k - \lambda\mathbf{A}^k\mathbf{D}^k](\mathbf{P}^{k\mathrm{T}}\mathbf{P}^k + \lambda\mathbf{D}^k)^{-1}. \tag{14}$$

[2]The popularity and advantages of the BCD approach to large-scale tensor decomposition were recently reviewed and highlighted in [19].

Algorithm 1: BTD-HIRLS-TC

Input: $\boldsymbol{\mathcal{Z}}$, λ, R_{ini}, L_{ini}
Output: Completed tensor, $\hat{\boldsymbol{\mathcal{Y}}}$
Initialize: $k=0$, $\mathbf{A}^0,\mathbf{B}^0,\mathbf{C}^0$
repeat
  Compute $\mathbf{D}_1^k,\mathbf{D}_2^k$ from (12) and (13)
  $\mathbf{D}^k \leftarrow (\mathbf{D}_1^k\otimes\mathbf{I}_L)\mathbf{D}_2^k$
  $\mathbf{P}^k \leftarrow \mathbf{B}^k\odot\mathbf{C}^k$
  Compute $\mathbf{A}^{k+1}$ from (14)
  $\mathbf{Q}^k \leftarrow \mathbf{C}^k\odot\mathbf{A}^k$
  Compute $\mathbf{B}^{k+1}$ from (15)
  $\mathbf{S}^k \leftarrow [\ (\mathbf{A}_1^k\odot_{\mathrm{c}}\mathbf{B}_1^k)\mathbf{1}_L \ \ \cdots \ \ (\mathbf{A}_R^k\odot_{\mathrm{c}}\mathbf{B}_R^k)\mathbf{1}_L\ ]$
  Compute $\mathbf{C}^{k+1}$ from (16)
  $k\leftarrow k+1$
until *convergence*
$\hat{\boldsymbol{\mathcal{Y}}} \leftarrow \sum_{r=1}^{R}\mathbf{A}_r^k\mathbf{B}_r^{k\mathrm{T}}\circ\mathbf{c}_r^k$

In analogy with (10), the objective function of the $\mathbf{B}$ sub-problem can be written as

$$f_{\mathbf{B}}(\mathbf{B}|\mathbf{A}^k,\mathbf{C}^k) = \tfrac{1}{2}\left\|\mathbf{Z}_{(2)}^{\mathrm{T}} - \mathbf{W}_{(2)}^{\mathrm{T}} * (\mathbf{Q}^k\mathbf{B}^{\mathrm{T}})\right\|_{\mathrm{F}}^2 + \lambda\sum_{r=1}^{R}\sqrt{\left(\sum_{l=1}^{L}\sqrt{\|\mathbf{a}_{rl}^k\|_2^2 + \|\mathbf{b}_{rl}\|_2^2+\eta^2}\right)^2 + \|\mathbf{c}_r^k\|_2^2+\eta^2},$$

which involves the mode-2 unfoldings. The majorization function $\mathbf{g_B}$ can be defined as above, replacing $\mathbf{P}$ by $\mathbf{Q}$. Its minimization has a unique solution, expressed as

$$\mathbf{B}^{k+1} = \mathbf{B}^k + [(\mathbf{Z}_{(2)} - \mathbf{W}_{(2)} * (\mathbf{B}^k\mathbf{Q}^{k\mathrm{T}}))\mathbf{Q}^k - \lambda\mathbf{B}^k\mathbf{D}^k](\mathbf{Q}^{k\mathrm{T}}\mathbf{Q}^k + \lambda\mathbf{D}^k)^{-1}. \tag{15}$$

The objective and majorization functions w.r.t. $\mathbf{C}$ can be analogously defined and shown to lead to the following closed-form expression for $\mathbf{C}^{k+1}$:

$$\mathbf{C}^{k+1} = \mathbf{C}^k + [(\mathbf{Z}_{(3)} - \mathbf{W}_{(3)} * (\mathbf{C}^k\mathbf{S}^{k\mathrm{T}}))\mathbf{S}^k - \lambda\mathbf{C}^k\mathbf{D}_1^k](\mathbf{S}^{k\mathrm{T}}\mathbf{S}^k + \lambda\mathbf{D}_1^k)^{-1}, \tag{16}$$

utilizing the mode-3 unfoldings of the tensors. As expected, instead of the composite matrix $\mathbf{D}$, only its $\mathbf{D}_1$ part appears in (16), promoting column sparsity for the factor $\mathbf{C}$.

The previous steps give rise to a novel BTD-based tensor completion algorithm, called BTD-HIRLS-TC and summarized in Algorithm 1.[3] Working as in [15], one can show that the basic iteration steps (14), (15), and (16) are in fact reweighted least squares recursions with a hierarchical structure, hence the name of the algorithm. BTD-HIRLS-TC is

[3]In fact, when $\boldsymbol{\mathcal{W}}$ is the all ones $I\times J\times K$ tensor, Algorithm 1 can be seen to reduce to the recently proposed BTD-HIRLS scheme [17, 15] for complete tensor rank-$(L_r,L_r,1)$ modeling.

based exclusively on computationally cheap matrix-wise operations, a rather uncommon feature of matrix/tensor completion algorithms, in which matrix factors are usually updated in a row-wise manner (e.g., [16]). This is due to the form of the approximate Hessian matrices (cf. (11)) adopted in the definitions of the quadratic upper bound functions and renders the algorithm suitable for large-scale completion problems. Moreover, the approximate Hessians satisfy the conditions of Assumption A in [20, Table 3] required for the BSUM convergence, and hence, by virtue of Theorem 1 and Example 1 of [20], every limit point of BTD-HIRLS-TC is a stationary point of the original objective function (9).

4. SIMULATION RESULTS

In this section, the performance of the proposed algorithm is evaluated with both synthetic and real data. Note that, by its construction, it admits a block-pruning mechanism, namely the $\mathbf{A}_r, \mathbf{B}_r$ blocks that correspond to columns of $\mathbf{C}$ with negligible energy may be removed as the algorithm progresses. This reduces the computational burden considerably while also providing estimates of R that improve in the course of the algorithm.

Consider first $60 \times 50 \times 55$ real-valued tensors $\boldsymbol{\mathcal{X}}$ that are generated as in (1) and are contaminated by additive noise, $\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{X}} + \sigma\boldsymbol{\mathcal{N}}$, where $\boldsymbol{\mathcal{N}}$ contains zero-mean, independent and identically distributed (i.i.d) Gaussian entries of unit variance and σ is set so that we get a signal-to-noise ratio (SNR), $\mathrm{SNR} = 10\log_{10}(\|\boldsymbol{\mathcal{X}}\|_{\mathrm{F}}^2/(\sigma^2\|\boldsymbol{\mathcal{N}}\|_{\mathrm{F}}^2))$, of 15 dB. The entries of the BTD factors are also sampled from a standard Gaussian distribution. R is set to 5 while the L_r's take values randomly in the range 2–9. The entries of $\boldsymbol{\mathcal{Y}}$ that are considered missing are uniformly sampled at random with a given rate. The algorithm is initialized with random factors and the *unknown* R and L_r's are all overestimated to 10. The relative recovery error, defined as $\mathrm{RE} = \|\hat{\boldsymbol{\mathcal{Y}}} - \boldsymbol{\mathcal{X}}\|_{\mathrm{F}}/\|\boldsymbol{\mathcal{X}}\|_{\mathrm{F}}$, is used as the figure of merit. Observe that this also measures the performance of denoising in addition to that of completion. For the stopping criterion, we use the relative difference between two consecutive REs. The algorithm stops either when this difference becomes less than 10^{-5} or a maximum of 200 iterations is reached. The regularization parameter λ is fine-tuned so that the minimum RE is attained. The median REs of 10 independent runs for each of three different missing rates are plotted in Fig. 1 versus time. Clearly, BTD-HIRLS-TC achieves a low RE in all cases, while requiring only a few seconds to converge. Its high computational efficiency should be attributed to its matrix-wise iteration steps that are to be contrasted with the row-wise updates in the corresponding algorithm of [16], requiring, in this experiment, more than 7 min to converge. Moreover, our experience with BTD-HIRLS-TC shows that it also estimates the model hyper-parameters R and L_r. Of course, it should be emphasized that the model selection part is here (compared to our earlier work [15]) more challenging given the incompleteness of the tensor. Nonetheless, the method's completion accuracy seems to be rather insensitive to rank over-estimation.

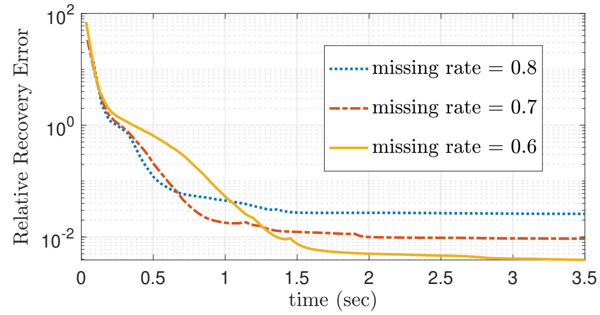
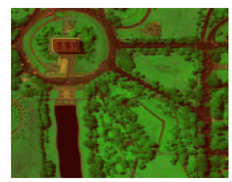
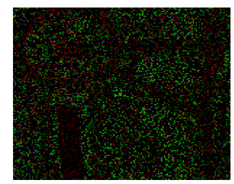


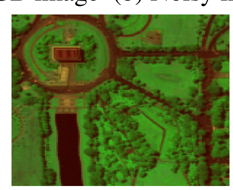
Fig. 1. RE of BTD-HIRLS-TC for different missing rates.



(a) Initial false RGB image



(b) Noisy incomplete image



(c) Reconstructed image

Fig. 2. Hyperspectral cube completion with BTD-HIRLS-TC.

We have also applied BTD-HIRLS-TC in the recovery of a highly incomplete real hyperspectral image (HSI). As demonstrated in, e.g., [21], the correlation structure of such imagery is well described by the BTD model, with R being the number of the end-members and the L_r's reflecting the ranks of the corresponding abundance maps. We consider the Washington DC Mall AVIRIS image captured at $K = 191$ contiguous spectral bands in the 0.4 to 2.4 μm region of the visible and infrared spectrum [22]. The size of the image is 150×150 pixels. Fig. 2 shows the RGB false color images reconstructed from bands (24,64,135) of (a) the original and (b) its noisy (with Gaussian i.i.d. noise added corresponding to SNR=15) and 80% incomplete HSI. The result of (denoising and) completion using BTD-HIRLS-TC, with R overestimated to 50 and all L_r's to 10, is shown in Fig. 2(c), clearly demonstrating the potential of the proposed scheme in HSI completion. In this experiment too, the proposed algorithm has exhibited its efficiency as compared to the corresponding scheme of [16] which took much more time to realize the task.

5. REFERENCES

[1] Q. Song, H. Ge, J. Caverlee, and X. Hu, "Tensor completion algorithms in big data analytics," *ACM Trans. Knowl. Discov. Data*, vol. 13, no. 1, Jan. 2019, Article 6.

[2] H. Huang, Y. Liu, Z. Long, and C. Zhu, "Robust low-rank tensor ring completion," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 1117–1126, July 2020.

[3] J. Wang, Y. Xia, and Y. Zhang, "Anomaly detection of hyperspectral image via tensor completion," *IEEE Geosci. Remote Sens. Lett.*, May 2020.

[4] M. A. Careem and A. Dutta, "Real-time prediction of non-stationary wireless channel," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 7836–7850, Dec. 2020.

[5] J. Luan and Z. Zhang, "Prediction of multidimensional spatial variation data via Bayesian tensor completion," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 39, no. 2, pp. 547–551, Feb. 2020.

[6] Y. A. Coutinho, N. Vervliet, L. De Lathauwer, and N. Moelans, "Combining thermodynamics with tensor completion techniques to enable multicomponent microstructure prediction," *npj Computational Materials*, vol. 6, no. 2, 2020.

[7] B. Yang, G. Wang, and N. D. Sidiropoulos, "Tensor completion via group-sparse regularization," in *Proc. ACSSC-2016*, Asilomar Conf. Grounds, Pacific Grove, CA, Nov. 2016.

[8] I. C. Tsaknakis, P. V. Giampouras, A. A. Rontogiannis, and K. D. Koutroumbas, "A computationally efficient tensor completion algorithm," *IEEE Signal Process. Lett.*, vol. 25, no. 8, pp. 1266–1270, Aug. 2018.

[9] M. Sørensen and L. De Lathauwer, "Fiber sampling approach to canonical polyadic decomposition and tensor completion," *SIAM J. Matrix Anal. Appl.*, vol. 40, no. 3, pp. 888—-917, 2019.

[10] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, and M. Akçakaya, "Tensor completion from regular sub-Nyquist samples," *IEEE Trans. Signal Process.*, vol. 68, pp. 1–16, 2020.

[11] X.-Y. Liu, S. Aeron, V. Aggarwal, and X. Wang, "Low-tubal-rank tensor completion using alternating minimization," *IEEE Trans. Inf. Theory*, vol. 66, no. 3, pp. 1714–1737, Mar. 2020.

[12] C.-Y. Ko, K. Batselier, L. Daniel, W. Yu, and N. Wong, "Fast and accurate tensor completion with total variation regularized tensor trains," *IEEE Trans. Image Process.*, vol. 29, pp. 6918–6931, May 2020.

[13] 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.

[14] L. De Lathauwer, "Decompositions of a higher-order tensor in block terms — Part II: Definitions and uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1033–1066, 2008.

[15] A. A. Rontogiannis, E. Kofidis, and P. V. Giampouras, "Block-term tensor decomposition: Model selection and computation," *IEEE J. Sel. Topics Signal Process.*, Jan. 2021.

[16] G. Zhang, X. Fu, J. Wang, X.-L. Zhao, and M. Hong, "Spectrum cartography via coupled block-term tensor decomposition," *IEEE Trans. Signal Process.*, vol. 68, pp. 3660–3675, May 2020.

[17] A. A. Rontogiannis, E. Kofidis, and P. V. Giampouras, "Block-term tensor decomposition: Model selection and computation," in *Proc. EUSIPCO-2020*, Amsterdam, The Netherlands, Aug. 2020.

[18] Y. Hu, C. Li, K. Meng, J. Qin, and X. Yang, "Group sparse optimization via $\ell_{p,q}$ regularization," *J. Mach. Learn. Res.*, vol. 18, no. 1, pp. 960–1011, 2017.

[19] X. Fu, N. Vervliet, L. De Lathauwer, K. Huang, and N. Gillis, "Computing large-scale matrix and tensor decomposition with structured factors — A unified nonconvex optimization perspective," *IEEE Signal Process. Mag.*, pp. 78–94, Sept. 2020.

[20] M. Hong, M. Razaviyayn, Z.-Q. Luo, and J.-S. Pang, "A unified algorithmic framework for block-structured optimization involving big data," *IEEE Signal Process. Mag.*, vol. 33, no. 1, pp. 57–77, Jan. 2016.

[21] Y. Qian, F. Xiong, S. Zeng, J. Zhou, and Y. Y. Tang, "Matrix-vector nonnegative tensor factorization for blind unmixing of hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 3, pp. 1776–1792, Mar. 2017.

[22] P. V. Giampouras, A. A. Rontogiannis, and K. D. Koutroumbas, "Alternating iteratively reweighted least squares minimization for low-rank matrix factorization," *IEEE Trans. Signal Process.*, vol. 67, no. 2, pp. 490–503, Jan. 2019.