BOOLEAN CP DECOMPOSITION OF BINARY TENSORS: UNIQUENESS AND ALGORITHM

Mamadou Diop[*,†], *Sebastian Miron*[*], *Antoine Souloumiac*[‡] *and David Brie*[*]

[*]Université de Lorraine, CRAN, CNRS, 54506 Vandoeuvre Cedex, France
[†]CEA Tech Grand-Est, 57075 Metz Technopôle Cedex 3, France
[‡]CEA Saclay, LIST, 91191 Gif-sur-Yvette Cedex, France
Email: *sebastian.miron@univ-lorraine.fr*

ABSTRACT

We propose an algorithm to perform the low-rank Boolean Canonical Polyadic Decomposition (BCPD) of a binary tensor. The proposed approach is based on the AO-ADMM strategy introduced in [1] and uses a post-nonlinear mixture model for binary sources. We show that this new method is better suited for low-rank approximation of binary tensors compared to other similar methods. We also provide an easy-to-check uniqueness condition for the BCPD. This is the first time that such a condition is derived for Boolean decompositions.

Index Terms— binary tensor, Boolean canonical polyadic decomposition, uniqueness, AO-ADMM

1. INTRODUCTION

Third-order *binary* tensors are three-way arrays with the entries composed of 0's and 1's. They are often used to capture ternary relationships, memberships or occurrences of events, *e.g.,* source IP - target IP - target port in network traffic analysis, adjacency matrices of a dynamic graph over time, predicate relations subject - object - verb in knowledge base, *etc.* To reveal latent structures in these binary tensors, the Boolean Canonical Polyadic Decomposition (BCPD) has been introduced; it allows to decompose a binary-valued tensor in a "logical sum" of rank-1 binary terms (sources) [2]. The BCPD is NP-hard and therefore suboptimal strategies have been proposed to tackle this difficulty. For example, in [3, 4] the *formal concept* analysis framework is used to achieve the decomposition, while in [2], the BCPD is obtained using an alternating approach based on the *discrete basis problem* for binary matrices, introduced in [5].

All these approaches are based on *greedy* strategies, designed to give optimal results in the case of exact decompositions and non-correlated sources. Their performance deteriorates rapidly in the presence of binary noise and therefore they are not well suited for low-rank decompositions of binary tensors. In this paper we propose an approach based on a relaxation of the BCPD problem over the nonnegative real orthant, coupled with a *post-nonlinear* model for the Boolean mixture of binary sources. We show that our approach yields very good results for low-rank approximation of binary tensors in the presence of sources having overlapping support (correlated sources). We also prove a sufficient condition for the uniqueness of the BCPD based on relationships between the support of the loading factors.

This work was supported by the operational program *FEDER-FSE Lorraine et Massif des Vosges 2014-2020.*

2. BOOLEAN CP DECOMPOSITION OF BINARY TENSORS

Consider a three-way binary data array (tensor) $\mathcal{X}$ of size $N \times M \times P$ such that its elements $\mathcal{X}_{nmp} \in \{0,1\}$ (with $n = 1,\ldots,N, m = 1,\ldots,M$ and $p = 1,\ldots,P$) can be expressed as:

$$\mathcal{X}_{nmp} = \bigvee_{k=1}^{K} (w_{nk} \wedge h_{mk} \wedge v_{pk}), \tag{1}$$

with $w_{nk}, h_{mk}, v_{pk} \in \{0,1\}$, and with "$\wedge$" and "$\vee$" denoting the "AND" and "OR" logical operators, respectively. Equation (1) expresses a third-order Boolean Canonical Polyadic Decomposition (BCPD) of rank K. As $w_{nk}, h_{mk}, v_{pk} \in \{0,1\}$, the logical "AND" operator in (1) can be equivalently replaced by the classical real numbers product, and thus $\mathcal{X}_{nmp} = \bigvee_{k=1}^{K} w_{nk}h_{mk}v_{pk}$. By regrouping the elements w_{nk}, h_{mk} and v_{pk} on the columns of matrices $\mathbf{W} = [\mathbf{w}_1 \ldots \mathbf{w}_K]$ $(N \times K)$, $\mathbf{H} = [\mathbf{h}_1 \ldots \mathbf{h}_K]$ $(M \times K)$ and $\mathbf{V} = [\mathbf{v}_1 \ldots \mathbf{v}_K]$ $(P \times K)$, respectively, the BCPD (1) can be expressed as:

$$\mathcal{X} = \bigvee_{k=1}^{K} (\mathbf{w}_k \circ \mathbf{h}_k \circ \mathbf{v}_k) = [\![\mathbf{W}, \mathbf{H}, \mathbf{V}]\!], \tag{2}$$

where "$\circ$" denotes the vector outer product and where the logical operation "$\vee$" is performed element-wise. Thus, in order to estimate the BCPD of $\mathcal{X}$, one must solve the following inverse problem:

$$\{\hat{\mathbf{W}}, \hat{\mathbf{H}}, \hat{\mathbf{V}}\} = \arg \min_{\mathbf{W},\mathbf{H},\mathbf{V}\in\{0,1\}} \left\| \mathcal{X} - \bigvee_{k=1}^{K} \mathbf{w}_k \circ \mathbf{h}_k \circ \mathbf{v}_k \right\|_F^2. \tag{3}$$

A classical way to perform the canonical polyadic decomposition in the real-valued case is to alternatingly estimate the three loading matrices using the three n-mode unfoldings of $\mathcal{X}$ [6]. A similar strategy can be used in the Boolean case, based on the three unfoldings hereafter:

$$\mathbf{X}_{(1)} = \mathbf{W} \diamond (\mathbf{V} \odot \mathbf{H})^T, \tag{4}$$

$$\mathbf{X}_{(2)} = \mathbf{H} \diamond (\mathbf{V} \odot \mathbf{W})^T, \tag{5}$$

$$\mathbf{X}_{(3)} = \mathbf{V} \diamond (\mathbf{H} \odot \mathbf{W})^T, \tag{6}$$

where "$\odot$" denotes the Khatri-Rao product and "$\diamond$" represents the *Boolean matrix product*, *i.e.*, the restriction of BCPD (2) to second-order tensors (matrices).

3. UNIQUENESS

The existing uniqueness conditions for the CP decomposition of real-valued tensors do not apply to the Boolean case. Therefore, before introducing the proposed approach for performing the BCPD, we analyze the uniqueness of this Boolean tensor decomposition. In [7] we derived a necessary and sufficient uniqueness condition for the Boolean decomposition of binary matrices. An extension of this condition to tensor case is possible but it does not have much practical interest because it would be very difficult to check. We give instead a sufficient condition for the uniqueness of the decomposition (2), much more easy to evaluate in practice. To our knowledge, this is the first time that such a condition is provided. We start by proving a sufficient condition for the Boolean decomposition of the binary matrix $\mathbf{X} = \mathbf{W} \diamond \mathbf{H} = \bigvee_{k=1}^{K} \mathbf{X}^{(k)} = \bigvee_{k=1}^{K} \mathbf{w}_k \mathbf{h}_k^T$.

Theorem 3.1 (Partial uniqueness of $\mathbf{X} = \mathbf{W} \diamond \mathbf{H}$)**.** *If* $supp^1(\mathbf{w}_\ell) \not\subseteq \bigcup_{k\neq\ell}^{K} supp(\mathbf{w}_k)$ *then the* ℓ^{th} *column of* $\mathbf{H}$, i.e, $\mathbf{h}_\ell$ *can be uniquely estimated from* $\mathbf{X}$. *(A similar condition can be proven for the uniqueness of* $\mathbf{w}_\ell$*).*

Proof. Suppose that $\mathbf{h}_\ell$ can not be uniquely estimated from $\mathbf{X}$, *i.e.* it $\exists\ \bar{\mathbf{X}}^{(\ell)} = \mathbf{w}_\ell \bar{\mathbf{h}}_\ell^T \neq \mathbf{X}^{(\ell)} = \mathbf{w}_\ell \mathbf{h}_\ell^T$ such that $\mathbf{X} = \bigvee_{k\neq\ell} \mathbf{w}_k \mathbf{h}_k^T \vee \mathbf{w}_\ell \mathbf{h}_\ell^T = \bigvee_{k\neq\ell} \mathbf{w}_k \mathbf{h}_k^T \vee \mathbf{w}_\ell \bar{\mathbf{h}}_\ell^T$. This is equivalent to $supp(\mathbf{X}) = \bigcup_{k=1}^{K} supp(\mathbf{w}_k \mathbf{h}_k^T) = \bigcup_{k\neq\ell} supp(\mathbf{w}_k \mathbf{h}_k^T) \cup supp(\mathbf{w}_\ell \mathbf{h}_\ell^T) = \bigcup_{k\neq\ell} supp(\mathbf{w}_k \mathbf{h}_k^T) \cup supp(\mathbf{w}_\ell \bar{\mathbf{h}}_\ell^T)$. Let us suppose, without loss of generality, that $\bar{\mathbf{h}}_\ell = \mathbf{h}_\ell \vee \mathbf{h} \neq \mathbf{h}_\ell$. Then, $supp(\mathbf{X}) = \bigcup_{k\neq\ell} supp(\mathbf{w}_k \mathbf{h}_k^T) \cup supp(\mathbf{w}_\ell (\mathbf{h}_\ell \vee \mathbf{h})^T) = \bigcup_{k\neq\ell} supp(\mathbf{w}_k \mathbf{h}_k^T) \cup supp(\mathbf{w}_\ell \mathbf{h}_\ell^T) \cup supp(\mathbf{w}_\ell \mathbf{h}^T)$. As $supp(\mathbf{h}) \not\subseteq supp(\mathbf{h}_\ell)$, it results that $supp(\mathbf{w}_\ell \mathbf{h}^T) \subseteq \bigcup_{k\neq\ell} supp(\mathbf{w}_k \mathbf{h}_k^T) \Leftrightarrow supp(\mathbf{w}_\ell) \subseteq \bigcup_{k\neq\ell} supp(\mathbf{w}_k)$ and $supp(\mathbf{h}) \subseteq \bigcup_{k\neq\ell} supp(\mathbf{h}_k)$, which ends the proof. □

[1]We define the support of a vector $\mathbf{x}$ as $supp(\mathbf{x}) = \{i, \mathbf{x}_i \neq 0\}$ and the support of matrix $\mathbf{X}$ as $supp(\mathbf{X}) = \{(i,j), \mathbf{X}_{ij} \neq 0\}$.

By applying theorem 3.1 to the unfoldings $\mathbf{X}_{(1)}, \mathbf{X}_{(2)}$ and $\mathbf{X}_{(3)}$ of $\mathcal{X}$, the following partial uniqueness condition for BCPD can be proven quite straightforwardly (for space reasons the proof will not be detailed in this version of the paper).

Theorem 3.2 (Partial uniqueness of the BCPD of $\mathcal{X}$)**.** *The* ℓ^{th} *rank-1 term* $\mathcal{X}^{(\ell)} = \mathbf{w}_\ell \circ \mathbf{h}_\ell \circ \mathbf{v}_\ell$ *in the BCPD* (2) *can be uniquely estimated from* $\mathcal{X}$ *if*

$$supp(\mathbf{w}_\ell) \not\subseteq \bigcup_{k\neq\ell}^{K} supp(\mathbf{w}_k) \text{ and } supp(\mathbf{h}_\ell) \not\subseteq \bigcup_{k\neq\ell}^{K} supp(\mathbf{h}_k)$$

or

$$supp(\mathbf{w}_\ell) \not\subseteq \bigcup_{k\neq\ell}^{K} supp(\mathbf{w}_k) \text{ and } supp(\mathbf{v}_\ell) \not\subseteq \bigcup_{k\neq\ell}^{K} supp(\mathbf{v}_k)$$

or

$$supp(\mathbf{h}_\ell) \not\subseteq \bigcup_{k\neq\ell}^{K} supp(\mathbf{h}_k) \text{ and } supp(\mathbf{v}_\ell) \not\subseteq \bigcup_{k\neq\ell}^{K} supp(\mathbf{v}_k).$$

If theorem 3.2 is satisfied for all values of $\ell = 1, \ldots, K$, we say that BPCD is *fully unique* or simply, *unique*.

4. PROPOSED APPROACH

We base our algorithm for solving the BCPD problem (3) on a *post-nonlinear* mixture model approach, similar to the one that we proposed in [7] for the matrix case. Instead of solving directly (3), we solve a relaxed version of it:

$$\{\hat{\mathbf{W}}, \hat{\mathbf{H}}, \hat{\mathbf{V}}\} = \arg \min_{\mathbf{W},\mathbf{H},\mathbf{V}\in\{0,1\}} \left\| \mathcal{X} - \Phi\left(\sum_{k=1}^{K} \mathbf{w}_k \circ \mathbf{h}_k \circ \mathbf{v}_k\right) \right\|_F^2, \tag{7}$$

where $\Phi(x)$ is the sigmoid function in Fig.1, applied element-wise.

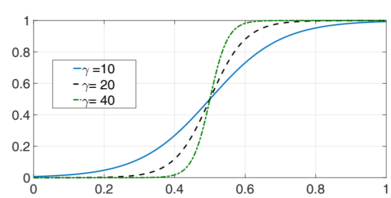

Fig. 1: Sigmoid function $\Phi(x) = \frac{1}{1+e^{-\gamma(x-0.5)}}$

In this paper we develop an algorithm for solving (7) inspired from the *Alternating Optimization - Alternating Direction Method of Multipliers (AO-ADMM)* introduced in [1]. The proposed approach can be summarized as follows:

Repeat

$$\min_{\mathbf{W},\bar{\mathbf{W}}} \frac{1}{2}\left\|\mathbf{X}_{(1)}-\Phi\left(\bar{\mathbf{W}}\left(\mathbf{V}\odot\mathbf{H}\right)^{T}\right)\right\|_F^2+\frac{\lambda}{2}\left\|\mathbf{W}-\mathbf{W}*\mathbf{W}\right\|_F^2$$
subject to $\mathbf{W}=\bar{\mathbf{W}}$

$$\min_{\mathbf{H},\bar{\mathbf{H}}} \frac{1}{2}\left\|\mathbf{X}_{(2)}-\Phi\left(\bar{\mathbf{H}}\left(\mathbf{V}\odot\mathbf{W}\right)^{T}\right)\right\|_F^2+\frac{\lambda}{2}\left\|\mathbf{H}-\mathbf{H}*\mathbf{H}\right\|_F^2$$
subject to $\mathbf{H}=\bar{\mathbf{H}}$

$$\min_{\mathbf{V},\bar{\mathbf{V}}} \frac{1}{2}\left\|\mathbf{X}_{(3)}-\Phi\left(\bar{\mathbf{V}}\left(\mathbf{H}\odot\mathbf{W}\right)^{T}\right)\right\|_F^2+\frac{\lambda}{2}\left\|\mathbf{V}-\mathbf{V}*\mathbf{V}\right\|_F^2$$
subject to $\mathbf{V}=\bar{\mathbf{V}}$

until convergence (8)

where "$*$" denotes the matrix Hadamard (element-wise) product. The second term in the expressions to minimize is used to constrain the entries of $\mathbf{W},\mathbf{H},\mathbf{W}$ to binarity, as explained in [7]. Using the results of [1], update rules can be obtained for the three minimization problems. For example, for the update of $\mathbf{W}$, the following expressions are obtained:

$$\bar{\mathbf{W}} \leftarrow \arg\min_{\bar{\mathbf{W}}} \frac{1}{2}\left\|\mathbf{X}_{(1)}-\Phi\left(\bar{\mathbf{W}}\left(\mathbf{V}\odot\mathbf{H}\right)^{T}\right)\right\|_F^2 + \frac{\rho}{2}\left\|\mathbf{W}-\bar{\mathbf{W}}+\mathbf{A}\right\|_F^2$$
$$\mathbf{W} \leftarrow \arg\min_{\mathbf{W}} \frac{\lambda}{2}\left\|\mathbf{W}-\mathbf{W}*\mathbf{W}\right\|_F^2+\frac{\rho}{2}\left\|\mathbf{W}-\bar{\mathbf{W}}+\mathbf{A}\right\|_F^2$$
$$\mathbf{A} \leftarrow \mathbf{A}+\mathbf{W}-\bar{\mathbf{W}},$$

with ρ a regularization parameter. The two minimization problems below are solved by gradient descent steps. The resulting algorithm, that we called *Boolean Tensor - ADMM (BT-ADMM)*, is summarized in **Algorithm 1.** Ω and Ψ are two matrix element-wise functions, that associate to each element $\mathbf{X}_{ij}$ of a matrix $\mathbf{X}$, the values $\frac{\gamma e^{-\gamma\cdot(\mathbf{X}_{ij}-0.5)}}{\left(1+e^{-\gamma\cdot(\mathbf{X}_{ij}-0.5)}\right)^2}$ and $\frac{\gamma e^{-\gamma\cdot(\mathbf{X}_{ij}-0.5)}}{\left(1+e^{-\gamma\cdot(\mathbf{X}_{ij}-0.5)}\right)^3}$, respectively. For the choice of the hyperparamters of the algorithm, an heuristic that works in most cases is the following: at iteration k choose the descent steps $\alpha^{(k)}$ proportional to $k^{-1/2}$, γ close to 1, $\lambda^{(k)} = 10\lambda^{(k-1)}$ (with $\lambda^{(0)}=10$) and $\rho = 50\lambda^{(0)}$.

5. RESULTS

In this section we illustrate the proposed approach in numerical simulations and compare it to similar methods of the state-of-the-art.

A first experiment illustrates the uniqueness condition for the BCPD. Figure 2 shows two rank-3 BCPD's of $10\times7\times5$ binary tensors (gray pixels symbolize the 1's). The first row of each figure represents the simulated data and the second

Algorithm 1 : BT-ADMM

Inputs: $\mathcal{X}$, K, Nb_{iter}, $Nb_{internal}$, ρ, λ, γ, ε, Nb_{grad1}, Nb_{grad2}
Outputs: $\mathbf{W}$, $\mathbf{H}$, $\mathbf{V}$
STEP 1: Initialization
  $\mathbf{W}\leftarrow$ rand(N,K), $\mathbf{H}\leftarrow$ rand(M,K), $\mathbf{V}\leftarrow$ rand(P,K)
  $\bar{\mathbf{W}}\leftarrow$ rand(N,K), $\bar{\mathbf{H}}\leftarrow$ rand(M,K), $\bar{\mathbf{V}}\leftarrow$ rand(P,K)
  $\mathbf{A}\leftarrow$ zeros(N,K), $\mathbf{B}\leftarrow$zeros(M,K), $\mathbf{C}\leftarrow$zeros(P,K)
STEP 2: Updates
for $t=1:Nb_{iter}$ **do**
  Update of $\bar{\mathbf{W}}$ and $\mathbf{W}$
  for $t_1=1:Nb_{internal}$ **do**
    for $t_{11}=1:Nb_{grad1}$ **do**
      $\bar{\mathbf{W}}\leftarrow\bar{\mathbf{W}}-\alpha_{\bar{\mathbf{W}}}(-\gamma(\mathbf{X}_{(1)}*\Omega(\bar{\mathbf{W}}(\mathbf{V}\odot\mathbf{H})^T))(\mathbf{V}\odot\mathbf{H})+\gamma\Psi(\bar{\mathbf{W}}(\mathbf{V}\odot\mathbf{H})^T)(\mathbf{V}\odot\mathbf{H})-\rho(\mathbf{W}-\bar{\mathbf{W}}+\mathbf{A}))$
    end for
    for $t_{12}=1:Nb_{grad2}$ **do**
      $\mathbf{W}\leftarrow\mathbf{W}-\alpha_{\mathbf{W}}(\lambda(\mathbf{W}-3\mathbf{W}^2+2\mathbf{W}^3)+\rho(\mathbf{W}-\bar{\mathbf{W}}+\mathbf{A}))$
    end for
    $\mathbf{A}=\mathbf{A}+\mathbf{W}-\bar{\mathbf{W}}$
  end for
  Update of $\bar{\mathbf{H}}$ and $\mathbf{H}$
  for $t_2=1:Nb_{internal}$ **do**
    for $t_{21}=1:Nb_{grad1}$ **do**
      $\bar{\mathbf{H}}\leftarrow\bar{\mathbf{H}}-\alpha_{\bar{\mathbf{H}}}(-\gamma(\mathbf{X}_{(2)}*\Omega(\bar{\mathbf{H}}(\mathbf{V}\odot\mathbf{W})^T))(\mathbf{V}\odot\mathbf{W})+\gamma\Psi(\bar{\mathbf{H}}(\mathbf{V}\odot\mathbf{W})^T)(\mathbf{V}\odot\mathbf{W})-\rho(\mathbf{H}-\bar{\mathbf{H}}+\mathbf{B}))$
    end for
    for $t_{22}=1:Nb_{grad2}$ **do**
      $\mathbf{H}\leftarrow\mathbf{H}-\alpha_{\mathbf{H}}(\lambda(\mathbf{H}-3\mathbf{H}^2+2\mathbf{H}^3)+\rho(\mathbf{H}-\bar{\mathbf{H}}+\mathbf{B}))$
    end for
    $\mathbf{B}=\mathbf{B}+\mathbf{H}-\bar{\mathbf{H}}$
  end for
  Update of $\bar{\mathbf{V}}$ and $\mathbf{V}$
  for $t_3=1:Nb_{internal}$ **do**
    for $t_{31}=1:Nb_{grad1}$ **do**
      $\bar{\mathbf{V}}\leftarrow\bar{\mathbf{V}}-\alpha_{\bar{\mathbf{V}}}(-\gamma(\mathbf{X}_{(3)}*\Omega(\bar{\mathbf{V}}(\mathbf{H}\odot\mathbf{W})^T))(\mathbf{H}\odot\mathbf{W})+\gamma\Psi(\bar{\mathbf{V}}(\mathbf{H}\odot\mathbf{W})^T)(\mathbf{H}\odot\mathbf{W})-\rho(\mathbf{V}-\bar{\mathbf{V}}+\mathbf{C}))$
    end for
    for $t_{32}=1:Nb_{grad2}$ **do**
      $\mathbf{V}\leftarrow\mathbf{V}-\alpha_{\mathbf{V}}(\lambda(\mathbf{V}-3\mathbf{V}^2+2\mathbf{V}^3)+\rho(\mathbf{V}-\bar{\mathbf{V}}+\mathbf{C}))$
    end for
    $\mathbf{C}=\mathbf{C}+\mathbf{V}-\bar{\mathbf{V}}$
  end for
  STEP 3: Stop criterion
  $\hat{\mathcal{X}}=\bigvee_{k=1}^{K}\mathbf{w}_k\circ\mathbf{h}_k\circ\mathbf{v}_k$
  if $\|\mathcal{X}-\hat{\mathcal{X}}\|^2+\sum_{i,k}\left(\mathbf{W}_{ik}^2-\mathbf{W}_{ik}\right)^2+\sum_{j,k}\left(\mathbf{H}_{jk}^2-\mathbf{H}_{jk}\right)^2+\sum_{\ell,k}\left(\mathbf{V}_{\ell k}^2-\mathbf{V}_{\ell k}\right)^2<\varepsilon$ **then**
    break
  end if
end for

row, the estimated BCPD. One can see that for Fig.2 (a) the uniqueness conditions of theorem 3.2 are verified for all 3 sources, while for Fig.2 (b) the partial uniqueness condition is not satisfied for the second source. Thus, for the second configuration, our algorithm yielded another decomposition that reproduces exactly the BCP model $\mathcal{X}$.

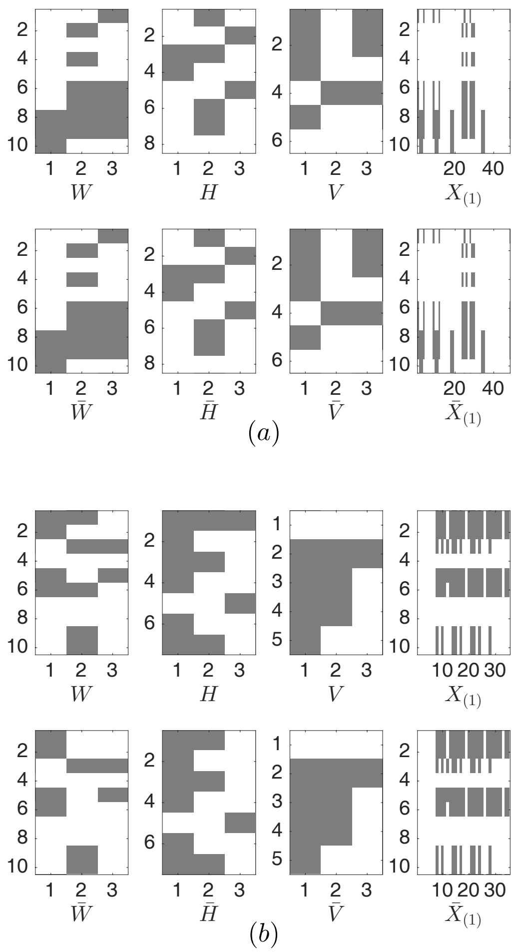

Fig. 2: Unique decomposition (a) and non-unique decomposition (b) of $\mathcal{X}$

The second experiment compares the performance of the proposed BT-ADMM algorithm to two other state-of-the-art approaches, the BCP-ALS of [2] and the T-FC of [4]. We plotted the estimation error for the loading matrices: $Error_{\mathbf{W},\mathbf{H},\mathbf{V}} = \frac{1}{3}\Big(\frac{\|\mathbf{W}-\hat{\mathbf{W}}\|_F^2}{NK} + \frac{\|\mathbf{H}-\hat{\mathbf{H}}\|_F^2}{MK} + \frac{\|\mathbf{V}-\hat{\mathbf{V}}\|_F^2}{PK}\Big)$ and the reconstruction error for $\mathcal{X}$: $Error_{\mathcal{X}} = \frac{\|\mathbf{X}_{(1)}-\hat{\mathbf{X}}_{(1)}\|_F^2}{NMP}$ versus the noise rate. For these simulations, we considered additive XOR noise generated according to a Bernoulli distribution of parameter b. The plotted points were averaged over 50 runs. Two scenarios were considered: in the first scenario (Fig.3) the sources were randomly simulated according to a Bernoulli distribution with parameter $p = 0.3$, *i.e.,* the sources have low correlation (their supports are approximately disjoint). In the second scenario (Fig. 4) we chose $p = 0.6$ in order to generate highly correlated sources.

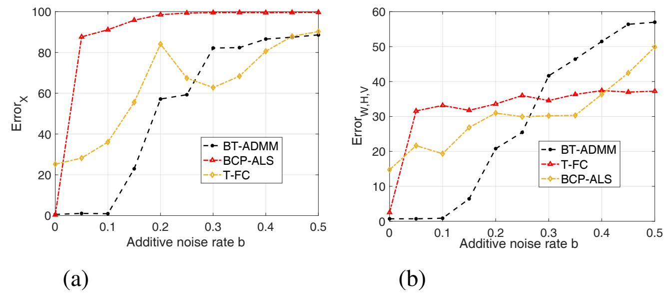

Fig. 3: (a) Reconstruction error for $\mathcal{X}$ and (b) estimation error for $\mathbf{W}, \mathbf{H}$ $\mathbf{V}$ vs. additive XOR noise rate b ($N = 20$, $M = 30$, $P = 10$, $K = 3$, $\rho = 10^9$) for $p = 0.3$

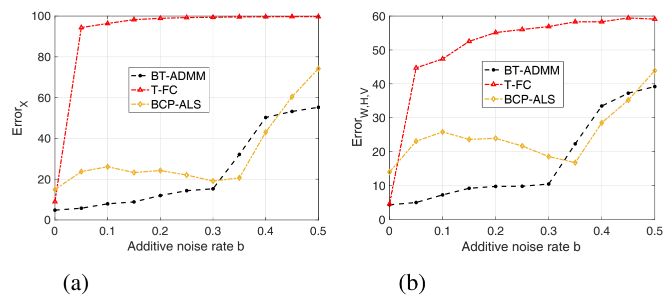

Fig. 4: (a) Reconstruction error for $\mathcal{X}$ and (b) estimation error for $\mathbf{W}, \mathbf{H}$ $\mathbf{V}$ vs. additive XOR noise rate b ($N = 20$, $M = 30$, $P = 10$, $K = 3$, $\rho = 10^9$) for $p = 0.6$

One can observe that for noise rates b inferior to 0.3 our algorithm yields better results than the competitor methods, which makes it an interesting tool for low-rank binary tensor approximation. For the values of $p > 0.3$ none of the three methods give good results because the XOR noise rate is high enough to completely destroy the low-rank structure of the data.

6. CONCLUSIONS

In this paper we introduced a new method for the Boolean canonical polyadic decomposition (BCPD) of binary-valued tensors based, on a post-nonlinear mixture model and an alternating ADMM approach. We illustrated in numerical simulations that our method outperforms similar state-of-the-art methods in the presence of XOR binary noise, which makes it well-adapted for low-rank binary tensor approximation. We also provided an easy-to-check sufficient condition for the uniqueness of BCPD; this is the first time that such a condition is derived.

7. REFERENCES

[1] Kejun Huang, Nicholas D. Sidiropoulos, and Athanasios P. Liavas, "A flexible and efficient algorithmic framework for constrained matrix and tensor factorization," *IEEE Transactions on Signal Processing*, vol. 64, no. 19, pp. 5052–5065, 2016.

[2] Pauli Miettinen, "Boolean tensor factorizations," in *IEEE 11th International Conference on Data Mining (ICDM 2011)*. IEEE, 2011, pp. 447–456.

[3] Radim Belohlavek and Vilem Vychodil, "Optimal factorization of three-way binary data," in *IEEE International Conference on Granular Computing (GrC 2010)*. IEEE, 2010, pp. 61–66.

[4] Radim Belohlavek, Cynthia Glodeanu, and Vilem Vychodil, "Optimal factorization of three-way binary data using triadic concepts," *Order*, vol. 30, no. 2, pp. 437–454, 2013.

[5] Pauli Miettinen, Taneli Mielikainen, Aristides Gionis, Gautam Das, and Heikki Mannila, "The discrete basis problem," *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 10, pp. 1348–1362, 2008.

[6] Tamara G. Kolda and Brett W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.

[7] Mamadou Diop, Anthony Larue, Sebastian Miron, and David Brie, "A post-nonlinear mixture model approach to binary matrix factorization," in *25th European Signal Processing Conference (EUSIPCO 2017)*. IEEE, 2017, pp. 321–325.