Coupled Block-term Tensor Decomposition Based Blind Spectrum Cartography

Guoyong Zhang
National Key Laboratory of Science and Technology on Communications
University of Electronic Science and Technology of China
Chengdu, China
15882214708@163.com

Xiao Fu
Electrical Engineering and Computer Science
Oregon State University
Corvallis, Oregon
xiao.fu@oregonstate.edu

Jun Wang
National Key Laboratory of Science and Technology on Communications
University of Electronic Science and Technology of China
Chengdu, China
junwang@uestc.edu.cn

Mingyi Hong
Electrical and Computer Engineering
University of Minnesota
Twin Cities, Minnesota
mhong@umn.edu

***Abstract*—Spectrum cartography aims at estimating the pattern of wideband signal power propagation over a region of interest (i.e. the radio map)—from limited samples taken sparsely over the region. Classical cartography methods are mostly concerned with recovering the aggregate radio frequency (RF) information while ignoring the constituents of the radio map—but fine-grained emitter-level RF information is of great interest. In addition, most existing cartography methods are based on random geographical sampling that is considered difficult to implement in some cases, due to legal/privacy/security issues. The theoretical aspects (e.g., identifiability of the radio map) of many existing methods are also unclear. In this work, we propose a radio map disaggregation method that is based on coupled block-term tensor decomposition. Our method guarantees identifiability of the individual wideband radio map of *each emitter* in the geographical region of interest (thereby that of the aggregate radio map as well), under some realistic conditions. The identifiability result holds under a large variety of geographical sampling patterns, including many pragmatic systematic sampling strategies. We also propose an effective optimization algorithm to carry out the formulated coupled tensor decomposition problem.**

***Index Terms*—coupled tensor decomposition, block term decomposition, radio map, spectrum cartography**

I. Introduction

Improving spectrum efficiency relies on accurate, fine-grained, and agile radio frequency (RF) awareness. Spectrum sensing is the first step towards this end. Spectrum sensing is commonly poised as a detection problem, which determines if a frequency band is used or not [1]; some works also consider it as a power spectral density (PSD) estimation problem that recovers a wideband PSD using a sub-Nyquist sampling rate [2]–[4]. RF cartography [5]–[8] takes one step further, which aims to construct a *spatial power loss field* (SLF) over a geographical area of interest. Because spectrum cartography provides valuable spatial information on top of spectral information, it is considered appealing for aggressive access, interference avoidance, and networking.

Spectrum sensing and cartography offer spectral occupancy and geographical loss field information to system designers. Nonetheless, RF awareness-enabling techniques are still far from being satisfactory. First, most of the aforementioned techniques only recover aggregate PSD or SLF that are normally contributed by multiple emitters. Simultaneously estimating PSD and SLF disaggregated to the emitter level has not been addressed. Second, most existing cartography techniques put emphases on the algorithmic aspects. However, it has been unclear if in theory the spectral and spatial information is recoverable, given limited samples over space and/or frequency. Third, most cartography methods work under random spatial sampling, while random sampling may not always be realistic, because some areas may be subject to security/privacy/legal constraints.

To handle the above challenges, in this work, we propose a new framework for spectrum cartography. Our method is based on a novel coupled tensor decomposition formulation, and admits a number of salient features. First, by modeling the wideband radio map as a tensor that follows the multilinear rank-$(L_r, L_r, 1)$ block-term decomposition model [9], our framework guarantees the identifiability of the SLF and PSD of each emitter, thereby achieving radio map disaggregation. Second, the proposed approach works under a number of systematic sampling strategies. This is particularly useful for sensing in urban areas, where sensors may not be deployed randomly. We also offer a simple alternating optimization algorithm to handle the proposed estimation criterion. A series of numerical simulations show that the proposed approach is effective and promising.

The work of X. Fu and M. Hong is supported in part by National Science Foundation under Project ECCS 1808159 and the Army Research Office under Project ARO W911NF-19-1-0247. G. Zhang and J. Wang are supported in part by the National Key R&D Program of China under Grant 2018YFC0807101, the National Research Program of China under Grant 9020302, the Foundation of National Key Laboratory of Science and Technology on Communications, the Innovation Fund of NCL (IFN), and the National Natural Science Foundation of China (NSFC) under Grant 61471099.

II. Signal Model and Problem Statement

We consider a scenario where R transmitters/sources exist within a 2-D geographical area of interest. Each source transmits over a certain frequency band which may overlap. Flat fading is assumed in the case where the frequency band of interest is relatively small compared to the carrier frequency, e.g., if the frequency bands span 20 MHz at a carrier frequency within 2-5 GHz. This way, the SLF of transmitter r is almost identical across different frequency bins $f_1, \ldots, f_K$. To be more precise, assuming that the signals from different transmitters are uncorrelated, the received aggregate PSD at 2-D geographical position (i, j) is represented as [5], [8]:

$$x_{i,j,k} = \sum_{r=1}^{R} \mathbf{S}_r(i,j) c_{k,r}, \tag{1}$$

where $\mathbf{S}_r \in \mathbb{R}^{I \times J}$ represents the (discretized) SLF of transmitter r , $c_{k,r}$ represents the PSD of source r measured at frequency bins f_k, and R the number of transmitters. The signal model (1) means that the received PSD is a superposition of each PSD scaled by its individual SLF.

The received space-space-frequency radio map can be naturally expressed in tensor notations. Define a tensor $\underline{\mathbf{X}} \in \mathbb{R}^{I \times J \times K}$ and a matrix $\mathbf{C} \in \mathbb{R}^{K \times R}$ such that

$$\underline{\mathbf{X}}(i,j,k) = x_{i,j,k}, \quad \mathbf{C}(k,r) = c_{k,r}.$$

Then, the signal model can be expressed as

$$\underline{\mathbf{X}} = \sum_{r=1}^{R} \mathbf{S}_r \circ \mathbf{c}_r, \tag{2}$$

where $\mathbf{c}_r$ denotes the rth column of $\mathbf{C}$ and $\circ$ denotes the outer product. We refer to this tensor $\underline{\mathbf{X}}$ as a *RF tensor*, since it reveals the RF environment across both the spatial and the spectral domains.

Accurately estimating $\underline{\mathbf{X}}$ is of great interest for many reasons—e.g., interference avoidance, routing, and source localization. Of course, $\underline{\mathbf{X}}$ cannot be fully observed. Many approaches have been proposed to estimate $\underline{\mathbf{X}}$ or 2-D versions of $\underline{\mathbf{X}}$ (i.e., SLF at a certain frequency band) from samples taken within the geographical area of interest [5]–[8]. This line of work is often referred to as *spectrum cartography*, which has drawn a lot attention. In this work, we take a step further: our goal is not only recovering the aggregate radio map from downsampled RF data, but also disaggregating $\underline{\mathbf{X}}$ to estimate SLF and PSD of each source. In other words, our goal is to estimate $\mathbf{S}_r$ and $\mathbf{c}_r$ for all r. This task is very well-motivated, since fine-grained emitter-level RF information can enable much more effective spectral use/reuse strategies.

III. Proposed Approach

A. BTD Model

Our approach starts from the following postulate: In (2), the SLF of a source r, denoted by $\mathbf{S}_r \in \mathbb{R}^{I \times J}$, is approximately a low-rank matrix. This postulate makes sense, since power propagation over space is continuous, which makes many columns and rows of $\mathbf{S}_r$ correlated. Consequently, each $\mathbf{S}_r$ can be approximated by a rank-L matrix, i.e.,

$$\mathbf{S}_r = \mathbf{A}_r \mathbf{B}_r^\top$$

with $L_r \ll \min\{I, J\}$. The low-rank assumption of $\mathbf{S}_r$ is reasonable, since SLF is smooth over spatial dimenson and thus has high correlation between adjacent columns/rows. Hence, we have

$$\underline{\mathbf{X}} = \sum_{r=1}^{R} \mathbf{S}_r \circ \mathbf{c}_r = \sum_{r=1}^{R} \left(\mathbf{A}_r \mathbf{B}_r^\top\right) \circ \mathbf{c}_r. \tag{3}$$

where $\mathbf{A}_r \in \mathbb{R}^{I \times L_r}$ and $\mathbf{B}_r \in \mathbb{R}^{J \times L_r}$ for $r = 1, \ldots, R$. We further denote

$$\mathbf{A} = [\mathbf{A}_1, \cdots, \mathbf{A}_R] \in \mathbb{R}^{I \times \sum_{r=1}^{R} L_r},$$
$$\mathbf{B} = [\mathbf{B}_1, \cdots, \mathbf{B}_R] \in \mathbb{R}^{J \times \sum_{r=1}^{R} L_r},$$
$$\mathbf{C} = [\mathbf{c}_1, \cdots, \mathbf{c}_R] \in \mathbb{R}^{K \times R}.$$

The key observation here is that the tensor $\underline{\mathbf{X}}$ in (3) follows a block-term decomposition (BTD) model [10] in multilinear rank-$(L_r, L_r, 1)$ terms. The BTD model has a very nice property—$\mathbf{A}_r \mathbf{B}_r^\top$ and $\mathbf{c}_r$ for $r = 1, \cdots, R$ are identifiable up to permutation and scaling ambiguities, under some mild conditions. The corresponding theorem is given as follows:

Theorem 1 *[9] Let $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ represent a BTD of $\underline{\mathbf{X}}$ in rank-$(L, L, 1)$ terms. Assume $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ are drawn from certain joint absolutely continuous distributions. If $IJ \geq L^2 R$ and*

$$\min\left(\left\lfloor \frac{I}{L} \right\rfloor, R\right) + \min\left(\left\lfloor \frac{J}{L} \right\rfloor, R\right) + \min(K, R) \geq 2R + 2,$$

then, $\{\mathbf{A}_r \mathbf{B}_r^\top, \mathbf{c}_r\}_{r=1}^{R}$ are essentially unique almost surely.

We briefly introduce the matricization of a BTD tensor. Some necessary definitions are required to support matricization. The classic Khatri-Rao product (column-wise Kronecker) is defined as

$$\mathbf{B}_r \odot \mathbf{A}_r = $$
$$[\mathbf{B}_r(:,1) \otimes \mathbf{A}_r(:,1), \cdots, \mathbf{B}_r(:,L) \otimes \mathbf{A}_r(:,L)] \in \mathbb{R}^{IJ \times L_r},$$

where $\otimes$ denotes the Kronecker product. The partionwise Khatri-Rao product between two partition matrices is defined as [9]

$$\mathbf{C} \odot_p \mathbf{A} = [\mathbf{c}_1 \otimes \mathbf{A}_1, \cdots, \mathbf{c}_R \otimes \mathbf{A}_R] \in \mathbb{R}^{IK \times \sum_{r=1}^{R} L_r}.$$

Using the above notations, the matricization of $\underline{\mathbf{X}} = \sum_{r=1}^{R} (\mathbf{A}_r \mathbf{B}_r^\top) \circ \mathbf{c}_r$ are expressed as follows:

$$\mathbf{X}_1 = (\mathbf{C} \odot_p \mathbf{B}) \mathbf{A}^\top, \tag{4a}$$
$$\mathbf{X}_2 = (\mathbf{C} \odot_p \mathbf{A}) \mathbf{B}^\top, \tag{4b}$$
$$\mathbf{X}_3 = [(\mathbf{B}_1 \odot \mathbf{A}_1)\mathbf{1}_{L_1}, \cdots, (\mathbf{B}_R \odot \mathbf{A}_R)\mathbf{1}_{L_R}]\mathbf{C}^\top,$$
$$= [\mathrm{vec}(\mathbf{A}_1 \mathbf{B}_1^\top), \cdots, \mathrm{vec}(\mathbf{A}_R \mathbf{B}_R^\top)]\mathbf{C}^\top, \tag{4c}$$

where the different matricizations are obtained by vectorizing different types of slabs (i.e., horizontal, vertical, and frontal slabs, respectively).

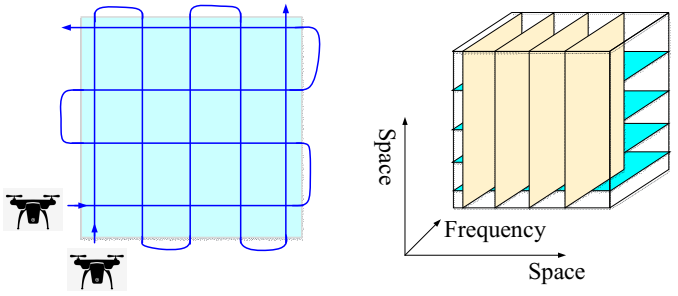

Fig. 1: Using moving sensors to sample the space-space-frequency RF tensor. (a) a possible sampling route; (b) the observed slabs (yellow and blue slabs).

B. Systematic Sampling

If $\underline{\mathbf{X}}$ is available, then estimating $\mathbf{S}_r$ and $\mathbf{c}_r$ can be done by directly applying BTD algorithms [10] to the tensor. In practice, however, the complete $\underline{\mathbf{X}}$ is not available. What we observe is a substantially undersampled version of $\underline{\mathbf{X}}$. In this paper, we design a regular sampling strategy where only some 'rows' and 'columns' are observed. The observed two subtensors are denoted by

$$\underline{\mathbf{X}}^{(1)} = \underline{\mathbf{X}}(\mathcal{S}_1, :, :), \ \underline{\mathbf{X}}^{(2)} = \underline{\mathbf{X}}(:, \mathcal{S}_2, :),$$

with $\mathcal{S}_1 \subset \{1, \cdots, I\}$ and $\mathcal{S}_2 \subset \{1, \cdots, J\}$. Define two row-selection matrices $\mathbf{P} = \mathbf{I}_I(\mathcal{S}_1, :)$ (select rows from an identity matrix by index set $\mathcal{S}_1$) and $\mathbf{Q} = \mathbf{I}_J(\mathcal{S}_2, :)$, and let $M = |\mathcal{S}_1|$ and $N = |\mathcal{S}_2|$, where $\mathbf{I}_I$ is an $I \times I$ identity matrix and $\mathbf{I}_J$ is a $J \times J$ identity matrix. Note that such a sampling pattern is not hard to implement. For example, a moving sensor (carried by unmaned vehicles or drones) can cruise over the region of interest and measure the PSD at each coordinate on its route, i.e., $\underline{\mathbf{X}}(i, j, :)$; see Fig.1 (a).

The sampled tensors can be represented as follows:

$$\underline{\mathbf{X}}^{(1)} = \sum_{r=1}^{R} ((\mathbf{P}\mathbf{A}_r)\mathbf{B}_r^\top) \circ \mathbf{c}_r,$$
$$\underline{\mathbf{X}}^{(2)} = \sum_{r=1}^{R} (\mathbf{A}_r(\mathbf{B}_r^\top \mathbf{Q}^\top)) \circ \mathbf{c}_r.$$

The sampled data is illustrated in Fig.1 (b). Our task now amounts to estimating $\mathbf{S}_r$ and $\mathbf{c}_r$ for $r = 1, \ldots, R$ from $\underline{\mathbf{X}}_1^{(1)}$ and $\underline{\mathbf{X}}_1^{(2)}$. We want to remark that the above sampling pattern is only one example under our framework. There are many more possible sampling patterns that can be derived.

C. Coupled Tensor Decomposition Approach

We aim to estimate $\mathbf{S}_r$ and $\mathbf{c}_r$ from $\underline{\mathbf{X}}^{(1)}$ and $\underline{\mathbf{X}}^{(2)}$ and propose the following criterion:

$$\min_{\{\mathbf{A}_r, \mathbf{B}_r\}_{r=1}^R, \mathbf{C}} \left\| \underline{\mathbf{X}}^{(1)} - \sum_{r=1}^{R} ((\mathbf{P}\mathbf{A}_r)\mathbf{B}_r^\top) \circ \mathbf{c}_r \right\|_F^2 + \left\| \underline{\mathbf{X}}^{(2)} - \sum_{r=1}^{R} (\mathbf{A}_r(\mathbf{B}_r^\top \mathbf{Q}^\top)) \circ \mathbf{c}_r \right\|_F^2 \quad (5)$$

We denote the objective function of Problem (5) as $J(\mathbf{A}, \mathbf{B}, \mathbf{C})$. This problem is not hard to handle. We propose a block coordinate descent (BCD) [11] based algorithm. To be specific, $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are updated in a cyclical fashion as follows:

$$\mathbf{A}^{t+1} \leftarrow \arg\min J(\mathbf{A}, \mathbf{B}^t, \mathbf{C}^t) \quad (6a)$$
$$\mathbf{B}^{t+1} \leftarrow \arg\min J(\mathbf{A}^{t+1}, \mathbf{B}, \mathbf{C}^t) \quad (6b)$$
$$\mathbf{C}^{t+1} \leftarrow \arg\min J(\mathbf{A}^{t+1}, \mathbf{B}^{t+1}, \mathbf{C}), \quad (6c)$$

where superscript t is the iteration index.

Note that each subproblem above is an unconstrained quadratic program—and thus is easy to solve. To proceed, take subproblem (6a) for example. Using the matricization based representation, solving (6a) is equivalent to solving the following:

$$\mathbf{A} \leftarrow \arg\min_{\mathbf{A}} \|\underline{\mathbf{X}}_1^{(1)} - (\mathbf{C} \odot_p \mathbf{B})(\mathbf{P}\mathbf{A})^\top\|_F^2 + \|\underline{\mathbf{X}}_1^{(2)} - (\mathbf{C} \odot_p \mathbf{Q}\mathbf{B})\mathbf{A}^\top\|_F^2.$$

The optimality condition of the above is as follows:

$$\underbrace{\mathbf{P}^\top\mathbf{P}}_{\mathbf{H}_1}\mathbf{A}\underbrace{(\mathbf{C} \odot_p \mathbf{B})^\top(\mathbf{C} \odot_p \mathbf{B})}_{\mathbf{H}_2} + \underbrace{\mathbf{I}_I}_{\mathbf{H}_3}\mathbf{A}\underbrace{(\mathbf{C} \odot_p \mathbf{Q}\mathbf{B})^\top(\mathbf{C} \odot_p \mathbf{Q}\mathbf{B})}_{\mathbf{H}_4} + = \underbrace{\mathbf{P}^\top\mathbf{X}_1^{(1)\top}(\mathbf{C} \odot_p \mathbf{B}) + \mathbf{X}_1^{(2)\top}(\mathbf{C} \odot_p \mathbf{Q}\mathbf{B})}_{\mathbf{H}_5}.$$

This simplifies the above equation as

$$\mathbf{H}_1\mathbf{A}\mathbf{H}_2 + \mathbf{H}_3\mathbf{A}\mathbf{H}_4 = \mathbf{H}_5.$$

To find $\mathbf{A}$ from the Sylvester equation $\mathbf{H}_1\mathbf{A}\mathbf{H}_2 + \mathbf{H}_3\mathbf{A}\mathbf{H}_4 = \mathbf{H}_5$, we propose to employ the extended Bartels-Stewart method [12].

The subproblems in (6b) and (6c) can be solved in the same way. In practice, one can also add regularization terms such as $\lambda\|\mathbf{A}\|_F^2$ for each latent factor to control scaling/counter-scaling issues, which will not complicate the algorithm.

D. Identifiability Analysis

In terms of identifiability, we show the following theorem:

Theorem 2 *Assume that $R \leq K$ and that $M \leq I$ and $N \leq J$. Also assume that $R + 2 \leq \min(\lfloor M/L \rfloor, R) + \min(\lfloor J/L \rfloor, R)$ and $N \geq L$, that $\mathbf{P}$ and $\mathbf{Q}$ both have full row-rank, and that $\mathbf{A}_r$, $\mathbf{B}_r$ and $\mathbf{C}$ are drawn from certain joint absolutely continuous distributions. Then, solving Problem* (5) *recovers $\{\mathbf{S}_r, \mathbf{c}_r\}$ for $r = 1 \ldots, R$ with probability one.*

Proof: Note that the optimal solution to Problem $(P1)$ should make the two terms zero. Let $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ denote the ground-truth and $(\widehat{\mathbf{A}}, \widehat{\mathbf{B}}, \widehat{\mathbf{C}})$ denote any optimal solution of Problem (5). We aim to prove that $\{\widehat{\mathbf{S}}_r, \widehat{\mathbf{c}}_r\}_{r=1}^R$ is essentially the ground-truth $\{\mathbf{S}_r, \mathbf{c}_r\}_{r=1}^R$ up to permutation and scaling ambiguities.

By Theorem 1, $\underline{\mathbf{X}}^{(1)}$ is identifiable. Then, $\{\mathbf{P}\mathbf{A}_r\mathbf{B}_r^\top\}_{r=1}^R$ and $\mathbf{C}$ can be identified up to scaling and permutation ambiguities. We have $\widehat{\mathbf{C}} = \mathbf{C}\mathbf{\Pi}\mathbf{\Lambda}$ ($\widehat{\mathbf{c}}_r = \lambda_r \mathbf{c}_{\pi_r}$), where $\mathbf{\Pi}$ is a permutation matrix and $\mathbf{\Lambda}$ is a scaling matrix. To be specific, $\mathbf{\Lambda} = \text{Diag}(\lambda_1, \cdots, \lambda_R)$ and $\pi_r \in \{1, \cdots, R\}$ satisfies $\mathbf{\Pi}(\pi_r, r) = 1$ for $r = 1, \cdots, R$. Note that

$$\begin{aligned}\underline{\mathbf{X}}_3^{(1)} &= [\text{vec}(\mathbf{P}\mathbf{A}_1\mathbf{B}_1^\top), \cdots, \text{vec}(\mathbf{P}\mathbf{A}_R\mathbf{B}_R^\top)]\mathbf{C}^\top \\ &= [\text{vec}(\mathbf{P}\widehat{\mathbf{A}}_1\widehat{\mathbf{B}}_1^\top), \cdots, \text{vec}(\mathbf{P}\widehat{\mathbf{A}}_R\widehat{\mathbf{B}}_R^\top)]\widehat{\mathbf{C}}^\top. \end{aligned} \quad (7)$$

Plugging $\widehat{\mathbf{C}} = \mathbf{C}\mathbf{\Pi}\mathbf{\Lambda}$ into (7), we have

$$\mathbf{P}\widehat{\mathbf{A}}_r\widehat{\mathbf{B}}_r^\top = \lambda_r^{-1}\mathbf{P}\mathbf{A}_{\pi_r}\mathbf{B}_{\pi_r}^\top.$$

Note that$\mathbf{A}_r, \widehat{\mathbf{A}}_r, \mathbf{B}_r, \widehat{\mathbf{B}}_r$ are all full column-rank matrices almost surely, and the row selection matrix $\mathbf{P}$ is also a full row-rank matrix. Hence, $\mathbf{P}\mathbf{A}_r \in \mathbb{R}^{M\times L}$ is a submatrix of $\mathbf{A}_r$ and $M \geq 2L$, which means that $\mathbf{P}\mathbf{A}_r$ is a full column-rank matrix and so is $\mathbf{P}\widehat{\mathbf{A}}_r$. Therefore, there exists an invertible matrix $\mathbf{F}_r \in \mathbb{R}^{L\times L}$ that satisfies $\widehat{\mathbf{B}}_r = \mathbf{B}_{\pi_r}\mathbf{F}_r$.

Considering the matricization of $\underline{\mathbf{X}}^{(2)}$, we have

$$\begin{aligned}\underline{\mathbf{X}}_3^{(2)} &= [\text{vec}(\mathbf{A}_1\mathbf{B}_1^\top\mathbf{Q}^\top), \cdots, \text{vec}(\mathbf{A}_R\mathbf{B}_R^\top\mathbf{Q}^\top)]\mathbf{C}^\top \\ &= [\text{vec}(\widehat{\mathbf{A}}_1\widehat{\mathbf{B}}_1^\top\mathbf{Q}^\top), \cdots, \text{vec}(\widehat{\mathbf{A}}_R\widehat{\mathbf{B}}_R^\top\mathbf{Q}^\top)]\widehat{\mathbf{C}}^\top. \end{aligned} \quad (8)$$

Plugging $\widehat{\mathbf{C}} = \mathbf{C}\mathbf{\Pi}\mathbf{\Lambda}$ into (8), we have

$$\widehat{\mathbf{A}}_r\widehat{\mathbf{B}}_r^\top\mathbf{Q}^\top = \lambda_r^{-1}\mathbf{A}_{\pi_r}\mathbf{B}_{\pi_r}^\top\mathbf{Q}^\top \quad (9)$$

Plugging $\widehat{\mathbf{B}}_r = \mathbf{B}_{\pi_r}\mathbf{F}_r$ into (9), we have

$$\widehat{\mathbf{A}}_r\mathbf{F}_r^\top\mathbf{B}_{\pi_r}^\top\mathbf{Q}^\top = \lambda_r^{-1}\mathbf{A}_{\pi_r}\mathbf{B}_{\pi_r}^\top\mathbf{Q}^\top.$$

Since $\mathbf{B}_{\pi_r}^\top\mathbf{Q}^\top$ is a full row rank matrix, we have $\widehat{\mathbf{A}}_r = \lambda_r^{-1}\mathbf{A}_{\pi_r}\mathbf{F}_r^{-\top}$. Therefore,

$$\begin{aligned}\widehat{\mathbf{S}}_r &= \widehat{\mathbf{A}}_r\widehat{\mathbf{B}}_r^\top = \lambda_r^{-1}\mathbf{A}_{\pi_r}\mathbf{F}_r^{-\top}(\mathbf{B}_{\pi_r}\mathbf{F}_r)^\top \\ &= \lambda_r^{-1}\mathbf{A}_{\pi_r}\mathbf{B}_{\pi_r}^\top = \lambda_r^{-1}\mathbf{S}_{\pi_r}.\end{aligned}$$

Therefore, we have proven that $\widehat{\mathbf{c}}_r = \lambda_r\mathbf{c}_{\pi_r}$ and $\widehat{\mathbf{S}}_r = \lambda_r^{-1}\mathbf{S}_{\pi_r}$ for $r = 1, \cdots, R$. □

Note that M and N denote the numbers of the sampled horizontal slabs and vertical slabs of the orginal RF tensor, respectively. This result is very encouraging because it gives the *first* identifiability-guaranteed formulation for RF tensor disaggregation using limited *regular* (i.e. non-random) samples. Perhaps a bit strikingly, Theorem 2 holds when the number of sampled slabs is larger than a certain threshold without assuming any randomness in sampling, without requiring how the samples should be taken (e.g., uniformly at random)—since this is reminiscent of identifiability of the BTD model instead of completion-based approaches. This is very appealing, since sampling strategies can be very flexible under our framework—and this can well accommodate scenarios where some regions cannot be sampled due to legal/privacy/security issues; see Fig. 2.

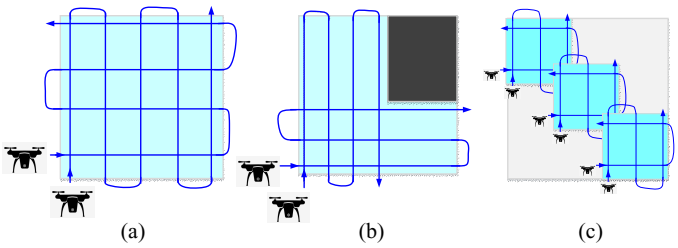

Fig. 2: Feasible sampling patterns, which can well accommodate scenarios where some regions cannot be sampled due to legal/privacy/security issues.

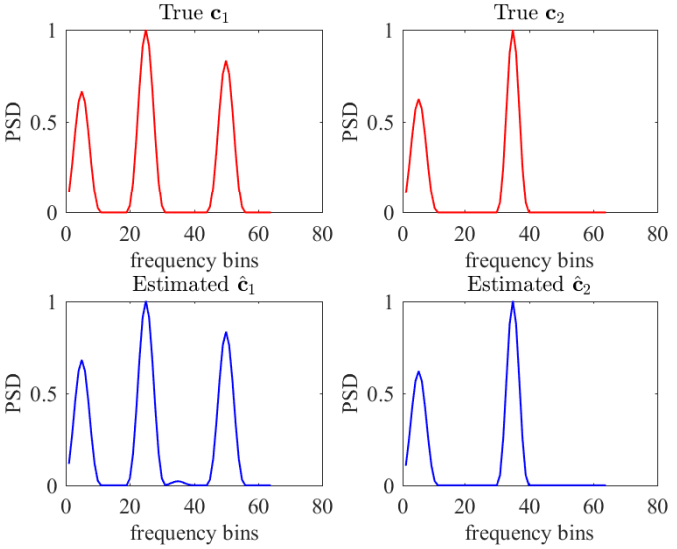

Fig. 3: Ground-truth (top) and estimated (bottom) PSDs of two sources.

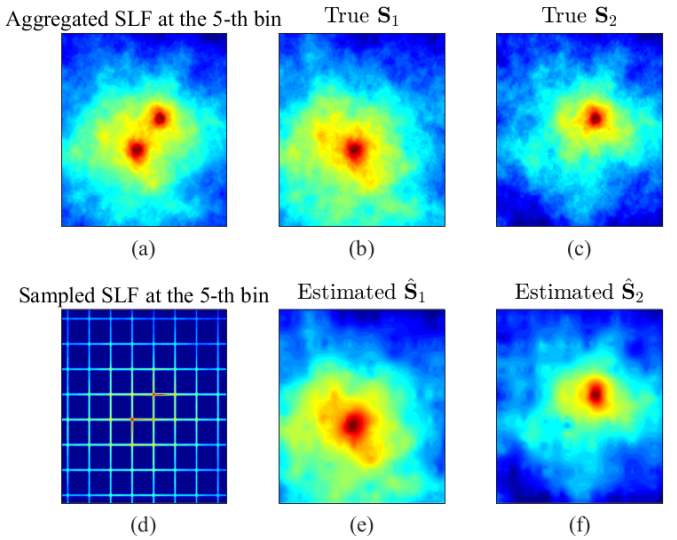

Fig. 4: (a) The aggregated SLF at the 5-th frequency bin, namely $\underline{\mathbf{X}}(:,:,5)$; (b) The ground-truth SLF of source 1; (c) The ground-truth SLF of source 1; (d) $M = 8$ rows and $N = 8$ columns are sampled from $\underline{\mathbf{X}}(:,:,5)$; (e) The estimated SLF of source 1; (f) The estimated SLF of source 2.

TABLE I: Performance under $L = 4$, $R = 3$, $N = 6$.

Algorithm	Measure	M			
		5	10	15	20
Proposed	$\text{NAE}_{\mathbf{C}}$	0.0381	0.0088	0.0055	0.0051
	$\text{NAE}_{\mathbf{S}}$	0.4717	0.3420	0.2474	0.1905
	$\text{NAE}_{\underline{\mathbf{X}}}$	0.3889	0.3017	0.2194	0.1693
	Running time(s)	1.1084	1.1716	1.4226	1.6292
SR	$\text{NAE}_{\mathbf{S}}$	0.4580	0.3208	0.2281	0.1884
	$\text{NAE}_{\underline{\mathbf{X}}}$	0.3795	0.2824	0.2083	0.1760
	Running time(s)	0.2883	0.4581	0.6992	0.9900
TPS	$\text{NAE}_{\mathbf{C}}$	0.5995	0.0246	0.0117	0.0110
	$\text{NAE}_{\mathbf{S}}$	0.9922	0.5430	0.3753	0.3249
	$\text{NAE}_{\underline{\mathbf{X}}}$	0.6671	0.4487	0.3226	0.2492
	Running time(s)	13.3200	21.3532	31.2258	43.0450
GLS	$\text{NAE}_{\mathbf{C}}$	0.5741	0.0216	0.0427	0.0251
	$\text{NAE}_{\mathbf{S}}$	0.9915	0.5518	0.4487	0.4402
	$\text{NAE}_{\underline{\mathbf{X}}}$	0.6549	0.4746	0.3196	0.2793
	Running time(s)	32.1612	59.2353	113.6869	195.5264

IV. Simulations

In this section, we present a number of simulations that showcase the effectiveness of our approach. The bands of interest are divided into $K = 64$ frequency bins. $R = 2$ transmitters (with unknown locations and PSDs) are randomly deployed in a region with a size of 100m $\times$ 100m. This region is uniformly discretized to $I \times J$ grids, where $I = J = 101$. Each SLF is generated by a path-loss model and spatial correlated log-normal shadowing modelas in [13], where have set the spatial decorrelation distance to be 30 m—which means a relatively severe shadowing effect. Note that in a typical outdoor scenario, the decorrelation distance ranges from 50 m to 100 m.

In the first simulation, $M = N = 8$ equi-spaced frontal and horizonal slabs are sampled, namely $\underline{\mathbf{X}}^{(1)} = \underline{\mathbf{X}}(\mathcal{S}_1, :, :)$ and $\underline{\mathbf{X}}^{(2)} = \underline{\mathbf{X}}(:, \mathcal{S}_2, :)$ are sampled (where $\mathcal{S}_1 = \mathcal{S}_2 = \{5, 18, 31, 44, 57, 70, 83, 96\}$). The ground-truth PSDs and estimated PSDs (after removing permutation and maximum amplitude be scaled to 1) of 2 sources are plotted in Fig. 3. The estimated PSDs and SFLs are shown in Fig. 4. One can see that the estimated PSDs and SFLs are almost visually identical to the ground truth, after some simple smoothing post-processing .

We also compare the performance with some benmarks, namely, the *semi-parametric regression* (SR) method [8], the TPS method [14], and the group lasso splines (GLS) method in [5]. Similar to our approach, the SR method estimates the SLF of each emitter, but it assumes that the PSDs are known. The GLS method and TPS recover the aggregated radio map. Hence, we use a BTD algorithm offered by the Tensorlab to disaggregate their estimated radio map. We use the normalized absolute error (NAE) w.r.t. ℓ_1 norm for $\mathbf{S}_r$, $\mathbf{c}_r$ and $\underline{\mathbf{X}}$ as our performance metric. We use the ℓ_1 norm since it is robust to outliers. All the simulations are run for 50 Monte Carlo trials and then taken the median of the NAEs.

Table I shows the results under $L = 4, R = 3, N = 6, K = 128$ and various M's. In general, all the algorithms favor the cases where M is large, since this means that more samples are available. SR is very efficient—since it does not need to estimate the PSDs. The PSD and SLF estimation accuracy of the proposed algorithm approaches that of SR, without knowing the PSDs. TPS and GLS work worse in estimating the PSDs and SLFs. This maybe because this approach did not utilize the BTD structure of the signal model.

V. Conclusion

In this work, we propose a coupled tensor decomposition framework for spectral cartography. Unlike existing works that mostly focus on estimating the aggregated radio map contributed by multiple emitters, our method can provably identify the SLFs and PSDs of the constituent emitters. In addition, our framework and identification theory work under a variety of systematic sampling strategies, instead of random sampling that is popular in the literature. This is particularly useful for sensing a region that is subject to legal/security/privacy regulations, e.g., an urban area, where random sensor deployment may not be possible.

References

[1] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 116–130, 2009.

[2] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio: State-of-the-art and recent advances," *IEEE Signal Process. Mag.*, vol. 29, no. 3, pp. 101–116, 2012.

[3] B. Hamdaoui, B. Khalfi, and M. Guizani, "Compressed wideband spectrum sensing: Concept, challenges, and enablers," *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 136–141, 2018.

[4] Z. Tian and G. B. Giannakis, "Compressed sensing for wideband cognitive radios," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2007, pp. 1357–1360.

[5] J. A. Bazerque, G. Mateos, and G. B. Giannakis, "Group-lasso on splines for spectrum cartography," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4648–4663, 2011.

[6] S.-J. Kim, E. Dall'Anese, and G. B. Giannakis, "Cooperative spectrum sensing for cognitive radios using kriged kalman filtering," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 1, pp. 24–36, 2011.

[7] J. A. Bazerque and G. B. Giannakis, "Distributed spectrum sensing for cognitive radio networks by exploiting sparsity," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1847–1862, 2010.

[8] D. Romero, S.-J. Kim, G. B. Giannakis, and R. López-Valcarce, "Learning power spectrum maps from quantized power measurements," *IEEE Trans. Signal Process.*, vol. 65, no. 10, pp. 2547–2560, 2017.

[9] 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.

[10] L. De Lathauwer and D. Nion, "Decompositions of a higher-order tensor in block terms—Part III: Alternating least squares algorithms," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1067–1083, 2008.

[11] D. P. Bertsekas, "Nonlinear programming," *Journal of the Operational Research Society*, vol. 48, no. 3, pp. 334–334, 1997.

[12] J. D. Gardiner, A. J. Laub, J. J. Amato, and C. B. Moler, "Solution of the sylvester matrix equation axb t+ cxd t= e," *ACM Transactions on Mathematical Software (TOMS)*, vol. 18, no. 2, pp. 223–231, 1992.

[13] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[14] S. Üreten, A. Yongaccoğlu, and E. Petriu, "A comparison of interference cartography generation techniques in cognitive radio networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2012, pp. 1879–1883.