A Characterization of Sampling Patterns for Low-Tucker-Rank Tensor Completion Problem

Morteza Ashraphijuo
Columbia University
Email: ashraphijuo@ee.columbia.edu

Vaneet Aggarwal
Purdue University
Email: vaneet@purdue.edu

Xiaodong Wang
Columbia University
Email: wangx@ee.columbia.edu

***Abstract*—In this paper, we characterize the deterministic conditions on the locations of the sampled entries, which are equivalent (necessary and sufficient) to finite completability of a tensor given some components of its Tucker rank. In order to derive this characterization, we propose an algebraic geometric analysis on the Tucker manifold, which allows us to incorporate multiple rank components in the proposed analysis in contrast with the conventional geometric approaches on the Grassmannian manifold. Then, using the developed tools for this analysis, we also derive a sufficient condition on the sampling pattern that ensures there exists only one completion for the sampled tensor (unique completability).**

I. Introduction

Most data around us are better represented with multiple dimensions to capture the correlations across different attributes and in many applications, part of the data may be missing. This paper investigates the fundamental conditions on the locations of the non-missing entries such that the multi-dimensional data can be recovered in finite and/or unique choices. In particular, we investigate deterministic conditions on the sampling pattern for finite or unique solution to a low-rank tensor completion problem given the sampled tensor and some of its Tucker rank components.

There are numerous applications of low-rank data completion in various areas including image or signal processing [1, 2], data mining [3], network coding [4], compressed sensing [5–7], reconstructing the visual data [8], and RF fingerprinting [9].

The majority of the literature on matrix and tensor completion are concerned with developing various optimization-based algorithms under some assumptions such as incoherence, etc., to construct a completion. In particular, low-rank matrix completion has been widely studied and many algorithms based on convex relaxation of rank [10–14] and alternating minimization [15], etc., have been proposed. For the tensor completion problem various solutions have been proposed that are based on convex relaxation of rank constraints [7, 16, 17], alternating minimization [8, 18].

Also, deterministic conditions on the sampling patterns have been studied for subspace clustering in [19–22]. In [23] by D. Pimentel-Alarcón *et. al.*, a deterministic sampling pattern is proposed that is necessary and sufficient for finite completability of the sampled matrix of the given rank. The same problem has been solved in [24] for a multi-view data, and also the tensor version of this problem given the canonical polyadic, Tucker and tensor-train ranks have been treated in [25], [26] and [27], respectively. Such an algorithm-independent condition can lead to a much lower sampling rate than that is required by the optimization-based completion algorithms. Moreover, in [28], deterministic conditions for multi-view data have been proposed. In this paper, we propose a geometric analysis on Tucker manifold to obtain deterministic conditions that lead to finite or unique completability for low Tucker rank tensors when multiple rank components are given.

This paper focuses on the low Tucker rank tensor completion problem, given a portion of the rank vector of the tensor. Specifically, we investigate the following two problems:

- **Problem (i):** Characterizing the necessary and sufficient conditions on the sampling pattern to have finitely many tensor completions for the given rank.
- **Problem (ii):** Characterizing conditions on the sampling pattern to ensure that there is exactly one completion for the given rank.

II. Preliminaries and Notations

In this paper, it is assumed that a d^{th}-order tensor $\mathcal{U} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ is sampled. For the sake of simplicity in notation, define $N \triangleq \left(\Pi_{j=1}^{d} n_j\right)$ and $N_{-i} \triangleq \frac{N}{n_i}$. Also, for any real number x, define $x^{+} \triangleq \max\{0, x\}$. Let $\mathbf{U}_{(i)} \in \mathbb{R}^{n_i \times N_{-i}}$ be the i-th matricization of the tensor $\mathcal{U}$ such that $\mathcal{U}(\vec{x}) = \mathbf{U}_{(i)}(x_i, \mathcal{M}_{(i)}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d))$, where $\mathcal{M}_{(i)}$ is an arbitrary bijective mapping $\mathcal{M}_{(i)} : (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d) \rightarrow \{1, 2, \ldots, N_{-i}\}$ and $\mathcal{U}(\vec{x})$ represents an entry of the tensor $\mathcal{U}$ with coordinate $\vec{x} = (x_1, \ldots, x_d)$.

Given $\mathcal{U} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ and $\mathbf{X} \in \mathbb{R}^{n_i \times n_i'}$, $\mathcal{U}' \triangleq \mathcal{U} \times_i \mathbf{X} \in \mathbb{R}^{n_1 \times \cdots \times n_{i-1} \times n_i' \times n_{i+1} \times \cdots \times n_d}$ is defined as

$$\mathcal{U}'(x_1, \cdots, x_{i-1}, k_i, x_{i+1}, \cdots, x_d) \triangleq \sum_{x_i=1}^{n_i} \mathcal{U}(x_1, \cdots, x_{i-1}, x_i, x_{i+1}, \cdots, x_d) \mathbf{X}(x_i, k_i). \quad (1)$$

Throughout this paper, we use Tucker rank as the rank of a tensor, which is defined as $\text{rank}(\mathcal{U}) = (r_1, \ldots, r_d)$ where $r_i = \text{rank}(\mathbf{U}_{(i)})$. The Tucker decomposition of a tensor $\mathcal{U}$ is given by

$$\mathcal{U} = \mathcal{C} \times_{i=1}^{d} \mathbf{T}_i, \quad (2)$$

978-1-5090-4096-4/17/$31.00 ©2017 IEEE

where $\mathcal{C} \in \mathbb{R}^{r_1 \times \cdots \times r_d}$ is the core tensor and $\mathbf{T}_i \in \mathbb{R}^{r_i \times n_i}$ are d orthogonal matrices. Then, (2) can be written as

$$\mathcal{U}(\vec{x}) = \sum_{k_1=1}^{r_1} \cdots \sum_{k_d=1}^{r_d} \mathcal{C}(k_1, \dots, k_d) \mathbf{T}_1(k_1, x_1) \dots \mathbf{T}_d(k_d, x_d).$$

The space of fixed Tucker-rank tensors is a manifold and the dimension of this manifold is shown in [29] to be $\sum_{i=1}^{d} \left(n_i \times r_i - r_i^2\right) + \Pi_{i=1}^{d}\ r_i$. Denote Ω as the binary sampling pattern tensor that is of the same size as $\mathcal{U}$ and $\Omega(\vec{x}) = 1$ if $\mathcal{U}(\vec{x})$ is observed and $\Omega(\vec{x}) = 0$ otherwise. For each subtensor $\mathcal{U}'$ of the tensor $\mathcal{U}$, define $N_\Omega(\mathcal{U}')$ as the number of observed entries in $\mathcal{U}'$ according to the sampling pattern Ω.

Problem Statement: We are interested in finding deterministic conditions on the sampling pattern tensor Ω under which there are infinite, finite, or unique completions of the sampled tensor $\mathcal{U}$ that satisfy rank$(\mathcal{U}) = (r_1, r_2, \dots, r_d)$.

III. Finite Completability

This section characterizes the connection between the sampling pattern and the number of solutions of a low-rank tensor completion.

A. Condition for Finite Completability Given Basis

Assume that the sampled tensor is $\mathcal{U} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ and rank components $\{r_{j+1}, \dots, r_d\}$ are given, where $j \in \{1, 2, \dots, d-1\}$ is an arbitrary fixed number. Without loss of generality assume that $r_{j+1} \geq \ldots \geq r_d$ throughout the paper.

Let the d^{th}-order tensor $\mathcal{V} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times r_{j+1} \times r_{j+2} \times \cdots \times r_d}$ be a basis of the sampled tensor $\mathcal{U} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$. Then, there exist full-rank matrices $\mathbf{T}_i$'s with $\mathbf{T}_i \in \mathbb{R}^{r_i \times n_i}$ such that

$$\mathcal{U} = \mathcal{V} \times_{i=j+1}^{d} \mathbf{T}_i, \tag{3}$$

or equivalently

$$\mathcal{U}(\vec{x}) = \sum_{k_{j+1}=1}^{r_{j+1}} \cdots \sum_{k_d=1}^{r_d} \mathcal{V}(x_1, \dots, x_j, k_{j+1}, \dots, k_d) \times \mathbf{T}_{j+1}(k_{j+1}, x_{j+1}) \dots \mathbf{T}_d(k_d, x_d). \tag{4}$$

For notational simplicity, define $\mathbb{T} = (\mathbf{T}_{j+1}, \dots, \mathbf{T}_d)$. Given $\mathcal{V}$, we are interested to find a subset of the mentioned polynomials that guarantees tuple $\mathbb{T}$ can be determined finitely.

Here, we briefly mention some key points to highlight the fundamentals of our proposed analysis.

- **Note** 1: As it can be seen from (4), any observed entry $\mathcal{U}(\vec{x})$ results in an equation that involves $\Pi_{i=j+1}^{d} r_i$ entries of $\mathcal{V}$ and also r_i entries of $\mathbf{T}_i$, $i = j+1, \dots, d$. Considering the entries of basis $\mathcal{V}$ and tuple $\mathbb{T}$ as variables (right-hand side of (4)), each observed entry results in a polynomial in terms of these variables.
- **Note** 2: For any observed entry $\mathcal{U}(\vec{x})$, the tuple $(x_1, \dots, x_j)$ specifies the coordinates of the $\Pi_{i=j+1}^{d} r_i$ entries of $\mathcal{V}$ that are involved in the corresponding polynomial.
- **Note** 3: For any observed entry $\mathcal{U}(\vec{x})$, the value of x_i specifies the column of the r_i entries of $\mathbf{T}_i$ that are involved in the corresponding polynomial, $i = j+1, \dots, d$.
- **Note** 4: Given all observed entries $\{\mathcal{U}(\vec{x}) : \Omega(\vec{x}) = 1\}$, we are interested in finding the number of possible solutions in terms of entries of $(\mathcal{V}, \mathbb{T})$ (infinite, finite or unique) via investigating the algebraic independence among these polynomials.
- **Note** 5: It can be concluded from Bernstein's theorem [30] that in a system of n polynomials in n variables that the coefficients are chosen generically, polynomials are algebraically independent with probability one, and therefore there exist only finitely many solutions. Moreover, in a system of n polynomials in $n-1$ variables (or less), polynomials are algebraically dependent with probability one.

The following definition will be used to determine the number of involved variables in a set of polynomials.

Definition 1. *For any $i \in \{j+1, \dots, d\}$ and $\mathcal{S}_i \subseteq \{1, \dots, n_i\}$, define $\mathcal{U}^{(\mathcal{S}_i)}$ as a set containing the entries of $|\mathcal{S}_i|$ rows (corresponding to the elements of $\mathcal{S}_i$) of $\mathbf{U}_{(i)}$. Moreover, define $\mathcal{U}^{(\mathcal{S}_{j+1}, \dots, \mathcal{S}_d)} = \mathcal{U}^{(\mathcal{S}_{j+1})} \cup \ldots \cup \mathcal{U}^{(\mathcal{S}_d)}$. Let τ be a subset of entries of $\mathcal{U}$. Then, $\mathcal{U}^{(\mathcal{S}_{j+1}, \dots, \mathcal{S}_d)}$ is called the minimal hull of τ if any $\mathcal{U}^{(\mathcal{S}_i)}$ includes exactly only those rows of $\mathbf{U}_{(i)}$ that include at least one of the entries in τ.*

Remark 1. Consider any set of polynomials $\{p_1, \dots, p_k\}$ in form of (4). Given the basis $\mathcal{V}$ these polynomials are in terms of entries of $\mathbb{T}$ and let τ be the set of corresponding entries to these polynomials in $\mathcal{U}$ and $\mathcal{U}^{(\mathcal{S}_{j+1}, \dots, \mathcal{S}_d)}$ be the minimal hull of τ. Let $\mathcal{S}$ denote the set of all variables (entries of $\mathbb{T}$) that are involved in at least one of the polynomials $\{p_1, \dots, p_k\}$. Recall that according to Note 3, if an entry of $\mathbf{T}_i$ is involved in a polynomial, all entries of the column that includes that entry are also involved in that polynomial. Therefore, $|\mathcal{S}| = \sum_{i=j+1}^{d} |\mathcal{S}_i| r_i$. $\square$

Note that each observed entry results in a scalar equation of the form of (4). Given the basis $\mathcal{V}$, we need at least $\sum_{i=j+1}^{d} (n_i r_i)$ polynomials to ensure the number of possible tuples $\mathbb{T}$ is not infinite since the number of variables (entries of $\mathbb{T}$) is $\sum_{i=j+1}^{d} (n_i r_i)$ in total. On the other hand, the $\sum_{i=j+1}^{d} (n_i r_i)$ mentioned polynomials should be algebraically independent to ensure the finiteness of tuples $\mathbb{T}$ since any algebraically independent polynomial reduces the dimension of the set of solutions by one. To ensure this independency, according to Note 5, any subset of t polynomials of the set of polynomials corresponding to the $\sum_{i=j+1}^{d} (n_i r_i)$ observed entries, should involve at least t variables. The following assumption will be used frequently as we show it satisfies the mentioned property.

Assumption A_j: Anywhere that this assumption is stated, there exist $\sum_{i=j+1}^{d} (n_i r_i)$ observed entries such that for any $\mathcal{S}_i \subseteq \{1, \dots, n_i\}$ for $i \in \{j+1, \dots, d\}$, $\mathcal{U}^{(\mathcal{S}_{j+1}, \dots, \mathcal{S}_d)}$ includes at most $\sum_{i=j+1}^{d} |\mathcal{S}_i| r_i$ of the mentioned $\sum_{i=j+1}^{d} (n_i r_i)$ observed entries.

Given the basis, the following lemma characterizes the necessary and sufficient condition on observed entries that leads to finite completability.

Lemma 1. *Assume that in* (3) *the basis tensor* $\mathcal{V} \in \mathbb{R}^{n_1\times\cdots\times n_j\times r_{j+1}\times\cdots\times r_d}$ *is given and* $\mathbf{T}_i \in \mathbb{R}^{r_i\times n_i}$ *are variables. Then, for almost every* $\mathcal{U}$ *with probability one, there are finitely many possible tuples* $\mathbb{T}$ *that satisfy* (3) *if and only if Assumption* A_j *holds.*

Proof. The $\sum_{i=j+1}^{d}(n_i r_i)$ observed entries results in $\sum_{i=j+1}^{d}(n_i r_i)$ scalar polynomials in terms of entries of $\mathbb{T}$ as in (3)-(4). We claim that any subset of these $\sum_{i=j+1}^{d}(n_i r_i)$ polynomials with t members involves at least t variables in total. Then, by Note 5, the proof of the lemma is complete. By contradiction, assume that there exists a subset of polynomials $\{p_1,\ldots,p_t\}$ that involves at most $t-1$ variables in total. Let τ be the subset of entries of $\mathcal{U}$ that result in polynomials $\{p_1,\ldots,p_t\}$ and denote the minimal hull of τ by $\mathcal{U}^{(\mathcal{S}_{j+1},\ldots,\mathcal{S}_d)}$. Observe that according to Remark 1, $\sum_{i=j+1}^{d}|\mathcal{S}_i| r_i \leq t-1$. On the other hand, Assumption A_j results that the number of polynomials in $\{p_1,\ldots,p_t\}$ is at most $\sum_{i=j+1}^{d}|\mathcal{S}_i| r_i$, i.e., $t \leq \sum_{i=j+1}^{d}|\mathcal{S}_i| r_i$. Hence, we have a contradiction, which completes the proof of the above claim. □

Remark 2. Assumption A_j results that given a basis there are finitely many tuples $\mathbb{T}$ such that (3) holds. Consequently, in what follows without loss of generality, we analyze the finite completability of basis $\mathcal{V}$ for one particular tuple $\mathbb{T}$ among all finitely many tuples. □

Since $\sum_{i=j+1}^{d}(n_i r_i)$ entries of the sampled tensor $\mathcal{U}$ are used to determine $\mathbb{T}$, in what follows we will use the polynomials corresponding to the set of the rest $N_\Omega(\mathcal{U})-\sum_{i=j+1}^{d}(n_i r_i)$ observed entries, denoted by $\mathcal{P}(\Omega)$, to obtain $\mathcal{V}$. Note that since $\mathbb{T}$ is already solved in terms of $\mathcal{V}$, each polynomial in $\mathcal{P}(\Omega)$ is in terms of elements of $\mathcal{V}$.

B. Constraint Tensor

<u>*Assumption* B_j</u>: $n_1 n_2 \ldots n_j \geq \sum_{i=j+1}^{d} r_i$.

For any subtensor $\mathcal{Y} \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_j\times 1\times\cdots\times 1}$ of the tensor $\mathcal{U}$, there exist row vectors $\theta_i \in \mathbb{R}^{r_i\times 1}$, $i=j+1,\ldots,d$, such that

$$\mathcal{Y}(x_1,\ldots,x_j,\vec{1}_{d-j}) = \sum_{k_{j+1}=1}^{r_{j+1}} \cdots \sum_{k_d=1}^{r_d} \mathcal{V}(x_1,\ldots,x_j,k_{j+1},\ldots,k_d)\theta_{j+1}(k_{j+1},1)\ldots\theta_d(k_d,1), \quad (5)$$

where $\vec{1}_{d-j}$ is an all-1 $(d-j)$-dimensional row vector.

For each subtensor $\mathcal{Y}$ of the sampled tensor $\mathcal{U}$, let $N_\Omega(\mathcal{Y}^{\mathbb{T}})$ denote the number of sampled entries in $\mathcal{Y}$ that have been used to obtain the tuple $\mathbb{T}$. Then, $\mathcal{Y}$ contributes $N_\Omega(\mathcal{Y})-N_\Omega(\mathcal{Y}^{\mathbb{T}})$ polynomial equations in terms of the entries of the basis $\mathcal{V}$ among all $N_\Omega(\mathcal{U})-\sum_{i=j+1}^{d}(n_i r_i)$ polynomials in $\mathcal{P}(\Omega)$.

The sampled tensor $\mathcal{U}$ includes $n_{j+1}n_{j+2}\cdots n_d$ subtensors that belong to $\mathbb{R}^{n_1\times n_2\times\cdots\times n_j\times 1\times\cdots\times 1}$ and we label these subtensors by $\mathcal{Y}_{(t_{j+1},\ldots,t_d)}$ where $(t_{j+1},\ldots,t_d)$ represents the coordinate of the subtensor. Define a binary valued tensor $\breve{\mathcal{Y}}_{(t_{j+1},\cdots,t_d)} \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_j\times\overbrace{1\times\ldots\times 1}^{d-j}\times k}$, where $k = N_\Omega(\mathcal{Y}_{(t_{j+1},\ldots,t_d)}) - N_\Omega(\mathcal{Y}^{\mathbb{T}}_{(t_{j+1},\ldots,t_d)})$ and its entries are described as the following. We can look at $\breve{\mathcal{Y}}_{(t_{j+1},\cdots,t_d)}$ as k tensors each belongs to $\mathbb{R}^{n_1\times n_2\times\cdots\times n_j\times 1\times\cdots\times 1}$. For each of the mentioned k tensors in $\breve{\mathcal{Y}}_{(t_{j+1},\cdots,t_d)}$ we set the entries corresponding to the $N_\Omega(\mathcal{Y}^{\mathbb{T}}_{(t_{j+1},\ldots,t_d)})$ observed entries that are used to obtain $\mathbb{T}$ in (3) equal to 1. For each of the other k observed entries, we pick one of the k tensors of $\breve{\mathcal{Y}}_{(t_{j+1},\cdots,t_d)}$ and set its corresponding entry (the same location as that specific observed entry) equal to 1 and set the rest of the entries equal to 0.

For the sake of simplicity in notation, we treat tensors $\breve{\mathcal{Y}}_{(t_{j+1},\cdots,t_d)}$ as a member of $\mathbb{R}^{n_1\times n_2\times\cdots\times n_j\times k}$ instead of $\mathbb{R}^{n_1\times n_2\times\cdots\times n_j\times\overbrace{1\times\cdots\times 1}^{d-j}\times k}$. Now, by putting together all $n_{j+1}n_{j+2}\cdots n_d$ tensors in dimension $(j+1)$, we construct a binary valued tensor $\breve{\mathbf{\Omega}} \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_j\times K_j}$, where $K_j = N_\Omega(\mathcal{U}) - \sum_{i=j+1}^{d}(n_i r_i)$ and call it the **constraint tensor**. In this paper, when we refer to subtensor of the constraint tensor that belongs to $\mathbb{R}^{n_1\times n_2\times\cdots\times n_j\times K'}$, it is assumed that all its subtensors belonging to $\mathbb{R}^{n_1\times n_2\times\cdots\times n_j\times 1}$ correspond to different subtensors of Ω that belongs to the same space.

C. Algebraic Independence

In this subsection, we derive the required number of algebraically independent polynomials in $\mathcal{P}(\Omega)$ for finite completability. Then, a sampling pattern on the constraint tensor is proposed to obtain the maximum number of algebraically independent polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}})$.

Lemma 2. *Assume that Assumptions* A_j *and* B_j *hold. For almost every* $\mathcal{U}$*, there exist only finitely many completions of* $\mathcal{U}$ *if and only if there exist* $\left(\Pi_{i=1}^{j} n_i\right)\left(\Pi_{i=j+1}^{d} r_i\right)-\left(\sum_{i=j+1}^{d} r_i^2\right)$ *algebraically independent polynomials in* $\mathcal{P}(\Omega)$.

Proof. Please refer to Lemma 2 in [26]. □

As a result of Lemma 2, we can certify finite completability based on the maximum number of algebraically independent polynomials in $\mathcal{P}(\Omega) = \mathcal{P}(\breve{\mathbf{\Omega}})$.

Definition 2. *Let* $\breve{\mathbf{\Omega}}' \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_j\times t}$ *be a subtensor of the constraint tensor* $\breve{\mathbf{\Omega}}$*. Let* $m_i(\breve{\mathbf{\Omega}}')$ *denote the number of nonzero columns of* $\breve{\mathbf{\Omega}}'_{(i)}$*. Also, let* $\mathcal{P}(\breve{\mathbf{\Omega}}')$ *denote the set of polynomials that correspond to nonzero entries of* $\breve{\mathbf{\Omega}}'$.

For any subtensor $\breve{\mathbf{\Omega}}' \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_j\times t}$ of the constraint tensor, the next theorem states an upper bound on the number of algebraically independent polynomials in the set $\mathcal{P}(\breve{\mathbf{\Omega}}')$. Recall that $\mathcal{P}(\breve{\mathbf{\Omega}}')$ includes exactly t polynomials. Define

$$g_{j+1}(x) = \sum_{i=j+1}^{d} \min\left\{ r_i, \left(x - \sum_{i'=j+1}^{i-1} r_{i'} \right)^{+} \right\} r_i. \quad (6)$$

Theorem 1. *Assume that Assumption B_j holds. For any subtensor $\breve{\mathbf{\Omega}}' \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times t}$ of the constraint tensor, the maximum number of algebraically independent polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}}')$ is no more than*

$$\left(\Pi_{i=j+1}^{d} r_i\right) m_{j+1}(\breve{\mathbf{\Omega}}') - g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}')), \tag{7}$$

where $g_{j+1}(\cdot)$ is given in (6).

Proof. Please refer to Theorem 1 in [26]. □

We are also interested in finding a condition on $\breve{\mathbf{\Omega}}'$ which results that $\mathcal{P}(\breve{\mathbf{\Omega}}')$ is minimally algebraically dependent, i.e., the polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}}')$ are algebraically dependent but polynomials in every of its proper subset are algebraically independent. This can help obtain the maximum number of algebraically independent polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}}')$ as Theorem 1 only provides an upper bound. The next lemma will be used in Theorem 2 in order to find a condition on $\breve{\mathbf{\Omega}}'$ which results that the set of polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}}')$ is minimally algebraically dependent.

Lemma 3. *Assume that Assumption B_j holds. Suppose that $\breve{\mathbf{\Omega}}' \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times t}$ is a subtensor of the constraint tensor such that $\mathcal{P}(\breve{\mathbf{\Omega}}')$ is minimally algebraically dependent. Then, for almost every $\mathcal{U}$, the number of variables that are involved in the set of polynomials $\mathcal{P}(\breve{\mathbf{\Omega}}')$ is $t-1$.*

Proof. Please refer to Lemma 3 in [26]. □

Theorem 2. *Assume that Assumption B_j holds. The polynomials in the set $\mathcal{P}(\breve{\mathbf{\Omega}})$ are algebraically dependent if and only if $\left(\Pi_{i=j+1}^{d} r_i\right) m_{j+1}(\breve{\mathbf{\Omega}}') - g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}')) < t$ for some subtensor $\breve{\mathbf{\Omega}}' \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times t}$ of the constraint tensor $\breve{\mathbf{\Omega}}$.*

Proof. If the polynomials in set $\mathcal{P}(\breve{\mathbf{\Omega}})$ are algebraically dependent, then there exists a subset of the polynomials that are minimally algebraically dependent. According to Lemma 3, if $\breve{\mathbf{\Omega}}' \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times t}$ is the corresponding subtensor to this minimally algebraically dependent set of polynomials, the number of variables that are involved in $\mathcal{P}(\breve{\mathbf{\Omega}}') = \{p_1, , p_2 \ldots, p_t\}$ is equal to $t-1$. On the other hand, $\left(\Pi_{i=j+1}^{d} r_i\right) m_{j+1}(\breve{\mathbf{\Omega}}') - g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}'))$ is the minimum possible number of involved variables in $\mathcal{P}(\breve{\mathbf{\Omega}}')$ (as we showed in the proof of Theorem 1) since $g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}'))$ is the maximum number of known entries of basis that are involved in $\mathcal{P}(\breve{\mathbf{\Omega}}')$. Therefore, $\left(\Pi_{i=j+1}^{d} r_i\right) m_{j+1}(\breve{\mathbf{\Omega}}') - g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}')) \leq t-1$.

In order to prove the other side of the statement, assume that $\left(\Pi_{i=j+1}^{d} r_i\right) m_{j+1}(\breve{\mathbf{\Omega}}') - g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}')) < t$ for some subtensor $\breve{\mathbf{\Omega}}' \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times t}$ of the constraint tensor $\breve{\mathbf{\Omega}}$. Recall that t is the number of polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}}')$. On the other hand, according to Theorem 1, $\left(\Pi_{i=j+1}^{d} r_i\right) m_{j+1}(\breve{\mathbf{\Omega}}') - g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}'))$ is the maximum number of algebraically independent polynomials, and therefore the polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}}')$ are not algebraically independent and it completes the proof. □

D. Finite Completability Using Analysis on Tucker Manifold

Theorem 2 together with Lemma 2 can lead to a necessary and sufficient condition on the constraint tensor $\breve{\mathbf{\Omega}}$ in order to ensure that there are finitely many completions for the sampled tensor $\mathcal{U}$, as stated by the next theorem.

Theorem 3. *Assume that Assumptions A_j and B_j hold. Then, for almost every $\mathcal{U}$, there are only finitely many tensors that fit in the sampled tensor $\mathcal{U}$, and have tensor rank components r_i for $i = j+1, \ldots, d$ if the following two conditions hold:*

(i) there exists a subtensor $\breve{\mathbf{\Omega}}' \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times n}$ of the constraint tensor such that $n = \left(\Pi_{i=1}^{j} n_i\right) \left(\Pi_{i=j+1}^{d} r_i\right) - \left(\sum_{i=j+1}^{d} r_i^2\right)$, and

(ii) for any $t \in \{1, \ldots, n\}$ and any subtensor $\breve{\mathbf{\Omega}}'' \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times t}$ of the tensor $\breve{\mathbf{\Omega}}'$ (in condition (i)), the following inequality holds

$$\left(\Pi_{i=j+1}^{d} r_i\right) m_{j+1}(\breve{\mathbf{\Omega}}'') - g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}'')) \geq t. \tag{8}$$

Proof. According to Lemma 2, for almost every $\mathcal{U}$, there are finitely many completions of $\mathcal{U}$ if and only if there exist $\left(\Pi_{i=1}^{j} n_i\right) \left(\Pi_{i=j+1}^{d} r_i\right) - \left(\sum_{i=j+1}^{d} r_i^2\right)$ algebraically independent polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}})$. On the other hand, according to Theorem 2 we conclude that a set of polynomials are algebraically independent if and only if condition (ii) in the statement of the theorem holds. Hence, for almost every $\mathcal{U}$, there are finitely many completions of $\mathcal{U}$ if and only if conditions (i) and (ii) hold. □

IV. Unique Completability

In this section, we provide the conditions on the sampling pattern to guarantee unique completability. The following assumptions is a stronger version of Assumption A_j to ensure that there exists only one tuple $\mathbb{T}$ given the basis.

Assumption A_j^+: Anywhere that this assumption is stated, there exist $\sum_{i=j+1}^{d} (n_i(r_i+1))$ observed entries such that for any $\mathcal{S}_i \subseteq \{1, \ldots, n_i\}$ for $i \in \{j+1, \ldots, d\}$, $\mathcal{U}^{(\mathcal{S}_{j+1}, \ldots, \mathcal{S}_d)}$ includes at most $\sum_{i=j+1}^{d} |\mathcal{S}_i|(r_i+1)$ of the mentioned $\sum_{i=j+1}^{d} (n_i(r_i+1))$ observed entries.

The following lemma gives the conditions under which a subset of entries of the basis $\mathcal{V}$ can be determined uniquely.

Lemma 4. *Assume that Assumption B_j holds. Suppose that $\breve{\mathbf{\Omega}}' \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times t}$ is a subtensor of the constraint tensor such that $\mathcal{P}(\breve{\mathbf{\Omega}}')$ is minimally algebraically dependent. Then, for almost every $\mathcal{U}$, all variables that are involved in the set of polynomials $\mathcal{P}(\breve{\mathbf{\Omega}}')$ can be uniquely determined.*

Proof. Please refer to Lemma 9 in [26]. □

Theorem 4. *Suppose that assumptions A_j^+ and B_j hold. Also, assume that there exist two disjoint subtensors $\breve{\mathbf{\Omega}}' \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times n}$ and $\breve{\mathbf{\Omega}}'_0 \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times n_0}$ of the constraint tensor such that $n = \left(\Pi_{i=1}^{j} n_i\right) \left(\Pi_{i=j+1}^{d} r_i\right) - \left(\sum_{i=j+1}^{d} r_i^2\right)$ and $n_0 = \left(\Pi_{i=1}^{j} n_i\right) - \lfloor \frac{\sum_{i=j+1}^{d} r_i^2}{\Pi_{i=j+1}^{d} r_i} \rfloor$ with the following conditions:*

(i) For any $t \in \{1,\dots,n\}$ and any subtensor $\breve{\mathbf{\Omega}}'' \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_j\times t}$ of the tensor $\breve{\mathbf{\Omega}}'$, the following inequality holds

$$\left(\Pi_{i=j+1}^{d} r_i\right) m_{j+1}(\breve{\mathbf{\Omega}}'') - g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}'')) \geq t. \qquad (9)$$

(ii) For any $t' \in \{1,\dots,n_0\}$ and any subtensor $\breve{\mathbf{\Omega}}''_0 \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_j\times t'}$ of the tensor $\breve{\mathbf{\Omega}}'_0$, the following inequality holds

$$\begin{aligned} & \left(\Pi_{i=j+1}^{d} r_i\right) m_{j+1}(\breve{\mathbf{\Omega}}''_0) - g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}''_0)) \\ \geq \quad & \left(\Pi_{i=j+1}^{d} r_i\right) t' - \left(\sum_{i=j+1}^{d} r_i^2\right)(t' - n_0 + 1)^+ . \quad (10) \end{aligned}$$

Then, for almost every $\mathcal{U}$ with probability one, there exists exactly one tensor that fits in the sampled tensor, and also rank$(\mathbf{U}_{(i)}) = r_i$, $i = j+1,\dots,d$. Therefore there is a unique completion of the sampled tensor with rank of $(r_1, r_2, \dots, r_d)$.

Proof. Please refer to Theorem 7 in [26]. □

V. Conclusions

This paper aims to find fundamental conditions on the sampling pattern for finite completability of a partially sampled tensor given its Tucker rank. To solve this problem, a novel geometric approach on Tucker manifold is proposed. Specifically, a set of polynomials based on the location of the sampled entries are first defined, and then a relationship between sampling patterns and the number of algebraically independent polynomials among all of the defined polynomials is characterized. Using this analysis, we have addressed two problems in this paper: (i) Characterizing the necessary and sufficient conditions on the sampling pattern to have finitely many tensor completions for the given Tucker rank, (ii) Characterizing conditions on the sampling pattern to ensure that there is exactly one completion for the given Tucker rank.

References

[1] E. J. Candès, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM review*, vol. 57, no. 2, pp. 225–251, 2015.

[2] H. Ji, C. Liu, Z. Shen, and Y. Xu, "Robust video denoising using low rank matrix completion." in *CVPR*. Citeseer, 2010, pp. 1791–1798.

[3] L. Eldén, *Matrix methods in data mining and pattern recognition*. SIAM, 2007, vol. 4.

[4] N. J. Harvey, D. R. Karger, and K. Murota, "Deterministic network coding by matrix completion," in *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, 2005, pp. 489–498.

[5] L.-H. Lim and P. Comon, "Multiarray signal processing: Tensor decomposition meets compressed sensing," *Comptes Rendus Mecanique*, vol. 338, no. 6, pp. 311–320, 2010.

[6] N. D. Sidiropoulos and A. Kyrillidis, "Multi-way compressed sensing for sparse low-rank tensors," *IEEE Signal Processing Letters*, vol. 19, no. 11, pp. 757–760, 2012.

[7] S. Gandy, B. Recht, and I. Yamada, "Tensor completion and low-n-rank tensor recovery via convex optimization," *Inverse Problems*, vol. 27, no. 2, pp. 1–19, 2011.

[8] X.-Y. Liu, S. Aeron, V. Aggarwal, and X. Wang, "Low-tubal-rank tensor completion using alternating minimization," *arXiv:1610.01690*, Oct. 2016.

[9] X.-Y. Liu, S. Aeron, V. Aggarwal, X. Wang, and M.-Y. Wu, "Tensor completion via adaptive sampling of tensor fibers: Application to efficient indoor RF fingerprinting," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016, pp. 2529–2533.

[10] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, 2009.

[11] M. Ashraphijuo, R. Madani, and J. Lavaei, "Characterization of rank-constrained feasibility problems via a finite number of convex programs," in *IEEE 55th Conference on Decision and Control (CDC)*, 2016, pp. 6544–6550.

[12] E. J. Candès and T. Tao, "The power of convex relaxation: Near-optimal matrix completion," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2053–2080, 2010.

[13] J. F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[14] M. Ashraphijuo, R. Madani, and J. Lavaei, "Inverse function theorem for polynomial equations using semidefinite programming," in *IEEE 54th Annual Conference on Decision and Control (CDC)*, 2015, pp. 6589–6596.

[15] P. Jain, P. Netrapalli, and S. Sanghavi, "Low-rank matrix completion using alternating minimization," in *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*. ACM, 2013, pp. 665–674.

[16] N. Kreimer, A. Stanton, and M. D. Sacchi, "Tensor completion based on nuclear norm minimization for 5D seismic data reconstruction," *Geophysics*, vol. 78, no. 6, pp. V273–V284, 2013.

[17] M. Signoretto, Q. T. Dinh, L. De Lathauwer, and J. A. Suykens, "Learning with tensors: a framework based on convex optimization and spectral regularization," *Machine Learning*, vol. 94, no. 3, pp. 303–351, 2014.

[18] X.-Y. Liu, S. Aeron, V. Aggarwal, and X. Wang, "Low-tubal-rank tensor completion using alternating minimization," in *SPIE Defense+ Security*. International Society for Optics and Photonics, 2016, pp. 984 809–984 809.

[19] D. Pimentel-Alarcón, L. Balzano, and R. Nowak, "Necessary and sufficient conditions for sketched subspace clustering," in *Allerton Conference on Communication, Control, and Computing*, 2016.

[20] D. Pimentel-Alarcón, N. Boston, and R. D. Nowak, "Deterministic conditions for subspace identifiability from incomplete sampling," in *IEEE International Symposium on Information Theory (ISIT)*, 2015, pp. 2191–2195.

[21] D. Pimentel-Alarcón, L. Balzano, R. Marcia, R. Nowak, and R. Willett, "Group-sparse subspace clustering with missing data," in *IEEE Workshop on Statistical Signal Processing (SSP)*, 2016, pp. 1–5.

[22] D. Pimentel-Alarcón, E. R. D. Nowak, and W. EDU, "The information-theoretic requirements of subspace clustering with missing data," in *International Conference on Machine Learning*, 2016.

[23] D. Pimentel-Alarcón, N. Boston, and R. Nowak, "A characterization of deterministic sampling patterns for low-rank matrix completion," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 4, pp. 623–636, 2016.

[24] M. Ashraphijuo, X. Wang, and V. Aggarwal, "Deterministic and probabilistic conditions for finite completability of low-rank multi-view data," *preprint arXiv:1701.00737*, Jan. 2017.

[25] M. Ashraphijuo and X. Wang, "Fundamental conditions for low-CP-rank tensor completion," *preprint arXiv:1703.10740*, 2017.

[26] M. Ashraphijuo, V. Aggarwal, and X. Wang, "Deterministic and probabilistic conditions for finite completability of low rank tensor," *preprint arXiv:1612.01597*, Dec. 2016.

[27] M. Ashraphijuo and X. Wang, "Characterization of deterministic and probabilistic sampling patterns for finite completability of low tensor-train rank tensor," *preprint arXiv:1703.07698*, 2017.

[28] M. Ashraphijuo, X. Wang, and V. Aggarwal, "A characterization of sampling patterns for low-rank multi-view data completion problem," in *IEEE International Symposium on Information Theory (ISIT)*, 2017.

[29] D. Kressner, M. Steinlechner, and B. Vandereycken, "Low-rank tensor completion by riemannian optimization," *BIT Numerical Mathematics*, vol. 54, no. 2, pp. 447–468, 2014.

[30] B. Sturmfels, *Solving Systems of Polynomial Equations*. American Mathematical Soc., 2002, no. 97.