Deterministic and Probabilistic Conditions for Finite Completability of Low-Tucker-Rank Tensor

Morteza Ashraphijuo, Vaneet Aggarwal, *Senior Member, IEEE*, and Xiaodong Wang, *Fellow, IEEE*

***Abstract*— We investigate the fundamental conditions on the sampling pattern, i.e., locations of the sampled entries, for finite completability of a low-rank tensor given some components of its Tucker rank. In order to find the deterministic necessary and sufficient conditions, 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. This analysis characterizes the algebraic independence of a set of polynomials defined based on the sampling pattern, which is closely related to finite completability of the sampled tensor, where finite completability simply means that the number of possible completions of the sampled tensor is finite. Probabilistic conditions are then studied and a lower bound on the sampling probability is given, which guarantees that the proposed deterministic conditions on the sampling patterns for finite completability hold with high probability. Furthermore, using the proposed geometric approach for finite completability, we propose a sufficient condition on the sampling pattern that ensures there exists exactly one completion of the sampled tensor.**

***Index Terms*— Low-rank tensor completion, finite completion, unique completion, Grassmannian manifold, Tucker manifold, extended Hall's theorem.**

I. Introduction

TENSORS are generalizations of vectors and matrices: a vector is a first-order tensor and a matrix is a second-order tensor. Most data around us are better represented with multiple dimensions to capture the correlations across different attributes. For example, a colored image can be considered as a third-order tensor, two of the dimensions (rows and columns) being spatial, and the third being spectral (color); while a colored video sequence can be considered as a fourth-order tensor, with time being the fourth dimension besides spatial and spectral dimensions. Similarly, a colored 3-D MRI image across time can be considered as a fifth-order tensor. 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 and probabilistic 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, i.e., ranks of some of its matricizations.

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], [9], seismic data processing [10]–[12], RF fingerprinting [13], [14], and reconstruction of cellular data [15].

The majority of the literature on matrix and tensor completion are concerned with developing various optimization-based algorithms under some assumptions such as incoherence [16], etc., to construct a completion. In particular, low-rank matrix completion has been widely studied and many algorithms based on convex relaxation of rank [17]–[19], non-convex optimization [20] and alternating minimization [16], etc., have been proposed. Also, a generalization of the low-rank matrix completion, which is completion from several low-rank sources has attracted attention recently [21]–[24]. For the tensor completion problem various solutions have been proposed that are based on convex relaxation of rank constraints [7], [10], [25]–[28], alternating minimization [9], [12], [29] and other heuristics [14], [30]–[32].

In the existing literature on optimization-based matrix or tensor completions, in addition to meeting the lower bound on the sampling probability, conditions such as incoherence [16], [32], [33], which constrains the values of the matrix or tensor entries, are required to obtain a completion with high probability. On the other hand, fundamental completability conditions that are independent of the specific completion algorithms have also been investigated. In [34]–[37] deterministic conditions on the locations of the sampled entries (sampling pattern) have been studied through algebraic geometry approaches on Grassmannian manifold that lead to finite and unique solutions to the matrix completion problem, where finite completability simply means that the number of possible completions of the sampled tensor is finite. Specifically, in [34] a deterministic sampling pattern is proposed that is necessary and sufficient for finite completability of the sampled matrix of the given rank. Such an algorithm-independent condition

Manuscript received December 5, 2016; revised May 2, 2019; accepted May 9, 2019. Date of publication May 28, 2019; date of current version August 16, 2019. This work was supported in part by the U.S. National Science Foundation (NSF) under Grant CCF-1814803, and in part by the U.S. Office of Naval Research (ONR) under Grant N000141410667. This work was presented in part at ISIT 2017, Aachen, Germany.

M. Ashraphijuo and X. Wang are with the Department of Electrical Engineering, Columbia University, Columbia, NY 10025 USA (e-mail: ashraphijuo@ee.columbia.edu; wangx@ee.columbia.edu).

V. Aggarwal is with the School of Industrial Engineering, Purdue University, West Lafayette, IN 47907 USA, and also with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA (e-mail: vaneet@purdue.edu).

Communicated by N. Kiyavash, Associate Editor for Statistical Learning.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2019.2919568

0018-9448 © 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

can lead to a much lower sampling rate than that is required by the optimization-based completion algorithms. For example, the required number of samples per column in [16] is on the order of $\mathcal{O}(\log(n)r^{2.5}\log(\|\mathbf{X}\|_F))$, where $\mathbf{X}$ is the unknown matrix with n rows and of rank r, while the required number of samples per column in [34] is on the order of $\mathcal{O}(\max\{\log(n), r\})$. The analysis on Grassmannian manifold in [34] is not capable of incorporating more than one rank constraint, and therefore this method is not efficient for solving the same problem for a tensor given multiple rank components. In this paper, we propose a geometric analysis on Tucker manifold to obtain deterministic and probabilistic conditions that lead to finite or unique completability for low-rank tensors when multiple rank components are given. Moreover, other related problems have been studied using algebraic geometry analysis, including high-rank matrix completion [38], rank estimation [39], and subspace clustering with missing data [24], [40]–[44].

This work is inspired by [34], where the analysis on Grassmannian manifold is proposed for a single-view matrix. Specifically, in [34] a novel approach is proposed to consider the rank factorization of a matrix and to treat each observed entry as a polynomial in terms of the entries of the components of the rank factorization. Then, under the genericity assumption, the algebraic independence among the mentioned polynomials is studied. In this paper, we consider the low-Tucker-rank tensor and follow the general approach that is similar to that in [34]. We mention some of the main differences: (i) geometry of the manifold, (ii) the equivalence class for the core tensor and consequently (iii) the canonical core tensor, (iv) structure of the polynomials, etc. are fundamentally different from those in [34]. Moreover, (v) the idea of using more than one rank constraint simultaneously in the algebraic geometry approach is also new. Hence, the manifold structure for the low-Tucker-rank tensor is fundamentally different from the Grassmannian manifold and we need to develop almost every step anew.

Tucker decomposition is a well-known method to represent a tensor [30], [45], [46]. In this paper, we use this decomposition to represent the sparsity of a tensor and use Tucker rank to model the low-rank structure of the tensor. There are several other well-known decompositions of a tensor as well, including polyadic decomposition [47], [48], tensor-train decomposition [49], [50], hierarchical Tucker representation [51], [52], tubal rank decomposition [53] and others.

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

- **Problem (i):** Characterizing the necessary and sufficient conditions on the sampling pattern to have finitely many tensor completions for the given rank.
  To solve this fundamental problem, we propose a geometric analysis framework on Tucker manifold. Specifically, we obtain a set of polynomials based on the location of the sampled entries in tensor and use Bernstein's theorem [54], [55] to identify the condition on the sampling pattern for ensuring sufficient number of algebraically independent polynomials in the mentioned set. Given any nonempty proper subset of the Tucker rank vector, this analysis leads to the necessary and sufficient condition on the sampling patterns for finite completability of the tensor. Given the entire Tucker rank vector this condition is sufficient for finite completability.
- **Problem (ii):** Characterizing conditions on the sampling pattern to ensure that there is exactly one completion for the given rank.
  We use our proposed geometric analysis for finite completability of low-rank tensors to obtain a sufficient conditions on the sampling patterns to ensure unique completability, which is milder than the sufficient condition for unique completability obtained through matricization analysis and applying the matrix method in [34].
- **Problem (iii):** If the elements in the tensor are sampled independently with probability p, what are the conditions on p such that the conditions in Problems (i) and (ii) are satisfied with high probability?
  We bound the number of needed samples to ensure the proposed sampling patterns for finite and unique tensor completability hold with high probability. Even though we follow a similar approach to [34] for the matrix case, we develop a generalization of Hall's theorem for bipartite graphs which is needed to prove the correctness of the bounds for both the tensor and the matrix cases. Moreover, it is seen that our proposed analysis on Tucker manifold leads to a much lower sampling rate than the corresponding analysis on Grassmannian manifold for both finite and unique tensor completions.

The remainder of this paper is organized as follows. In Section II, some preliminaries and notations are presented, and also an example is given that illustrates the advantage of tensor analysis over analyzing matricizations of a tensor. In Section III, Problem (i) is studied and the sampling patterns that ensure finite completions are found using tensor algebra. In Section IV, we study Problem (iii) for the case of finite completion and the key to solving this problem is the proof of the generalized Hall's theorem, which is an independent result in graph theory. Section V considers Problem (ii) to give a sufficient condition on the sampling pattern for unique completability. Further, Problem (iii) for unique completion is also studied. Some numerical results are provided in Section VI to compare the sampling rates for finite and unique completions based on our proposed tensor analysis versus the matricization method. Finally, Section VII concludes the paper.

II. Background

A. Preliminaries and Notations

In this paper, it is assumed that a dth-order tensor $\mathcal{U} \in \mathbb{R}^{n_1\times\cdots\times n_d}$ is chosen generically from the manifold of $n_1 \times \cdots \times n_d$ tensors of the given Tucker rank (will be explained rigorously later). 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

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

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

$$\mathcal{U}(\vec{x}) = \sum_{k_1=1}^{r_1} \cdots \sum_{k_d=1}^{r_d} \mathcal{C}(k_1, \ldots, k_d)\mathbf{T}_1(k_1, x_1) \ldots \mathbf{T}_d(k_d, x_d). \quad (3)$$

The space of fixed Tucker-rank tensors is a manifold and the dimension of this manifold is shown in [30] 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.

B. Problem Statement and a Motivating Example

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 $\text{rank}(\mathcal{U}) = (r_1, r_2, \ldots, r_d)$. Moreover, we are interested in finding probabilistic sampling strategies that ensure the obtained conditions for finite and unique completability hold, respectively, with high probability. The matrix version of this problem has been treated in [34]. In this paper, we investigate this problem for general order tensors.

In this subsection, we intend to compare the following two approaches in an example to emphasize the necessity of our analysis for general order tensors: (i) analyzing each matricization individually with the rank constraint of the corresponding matricization, (ii) analyzing via Tucker decomposition. In particular, we will show via an example that analyzing each of the matricizations separately is not enough to guarantee finite completability when multiple rank components are given. On the other hand, we show that for the same example Tucker decomposition ensures finite completability. Hence, this example illustrates that matricization analysis does not take advantage of the full information of given Tucker rank and thus fails to provide a necessary and sufficient condition for finite completability when more than one component of the rank vector is given.

Consider a 3rd-order tensor $\mathcal{U} \in \mathbb{R}^{2\times2\times2}$ with Tucker rank $(1, 1, 1)$. First, we show that having any 4 entries of $\mathcal{U}$, there are infinitely many completions of any matricization with the corresponding rank constraint. Hence, along each dimension there exist a set of infinite completions given the corresponding rank constraint. Note that the analysis on Grassmannian manifold in [34] is not capable of incorporating more than one rank constraint. However, as we show it is possible that the intersection of the mentioned three infinite sets is a finite set and that is why we need an analysis that is able to incorporate more than one rank constraint. Without loss of generality, it suffices to show the claim only for its first matricization. Therefore, the claim reduces to the following statement:

Statement: *Having any* 4 *entries of a rank-*1 *matrix* $\mathbf{U} \in \mathbb{R}^{2\times4}$*, there are infinitely many completions for it.*

In order to prove the above statement, we need to consider the following four possible scenarios:

(i) The 4 observed entries are in a row. In this case, clearly, there are infinitely many completions for the other row as it can be any scalar multiplied by the first row.

(ii) The 4 observed entries are such that there is a column in which there is no observed entries. In this case, there are infinitely many completions for this column as it can be any scalar multiplied by the other columns.

(iii) The 4 observed entries are such that there is one observed entry in each column, and also each row has exactly two observed entries. Assume that the two observed entries in the second row are the pair (a, b). In this case, for every pair (ka, kb) as the value of the two non-observed entries of the first row (where k is an arbitrary scalar) there is a unique completion for the rest of the entries. As a result, there are infinitely many completions for this matrix.

(iv) The 4 observed entries are such that there is one observed entry in each column, and also the first and second rows have 3 and 1 observed entries, respectively. In this case, for each value of the only non-observed entry of the first row there is a unique completion. Therefore, there are infinitely many completions for this matrix.

Assume that the entries $\mathcal{U}(1, 1, 1)$, $\mathcal{U}(2, 1, 1)$, $\mathcal{U}(1, 2, 1)$, and $\mathcal{U}(1, 1, 2)$ are observed. Now, we take advantage of all elements of Tucker rank simultaneously, in order to show there are only finitely many tensor completions. Using Tucker decomposition (2), and given the rank is $(1, 1, 1)$, without loss of generality, assume that the scalar $\mathcal{C} = 1$ and $\mathbf{T}_1 = (x, x')$, $\mathbf{T}_2 = (y, y')$ and $\mathbf{T}_3 = (z, z')$, and then the following equalities hold

$$\begin{aligned} \mathcal{U}(1,1,1) &= xyz, & \mathcal{U}(2,2,1) &= x'y'z, \\ \mathcal{U}(2,1,1) &= x'yz, & \mathcal{U}(2,1,2) &= x'yz', \\ \mathcal{U}(1,2,1) &= xy'z, & \mathcal{U}(1,2,2) &= xy'z', \\ \mathcal{U}(1,1,2) &= xyz', & \mathcal{U}(2,2,2) &= x'y'z'. \end{aligned} \quad (4)$$

The unknown entries can be determined uniquely in terms of the 4 observed entires as

$$\mathcal{U}(2,2,1) = x'y'z = \frac{\mathcal{U}(2,1,1)\mathcal{U}(1,2,1)}{\mathcal{U}(1,1,1)},$$
$$\mathcal{U}(2,1,2) = x'yz' = \frac{\mathcal{U}(2,1,1)\mathcal{U}(1,1,2)}{\mathcal{U}(1,1,1)},$$
$$\mathcal{U}(1,2,2) = xy'z' = \frac{\mathcal{U}(1,2,1)\mathcal{U}(1,1,2)}{\mathcal{U}(1,1,1)},$$
$$\mathcal{U}(2,2,2) = x'y'z' = \frac{\mathcal{U}(2,1,1)\mathcal{U}(1,2,1)\mathcal{U}(1,1,2)}{\mathcal{U}(1,1,1)\mathcal{U}(1,1,1)}. \quad (5)$$

Therefore, considering the Tucker decomposition, there is only one (finite) completion(s) having this particular 4 observed entries as above. Note that only given $r_2 = r_3 = 1$, it can be verified using Tucker decomposition similarly that the completion is still unique.

III. Deterministic Conditions for Finite Completability

This section characterizes the connection between the sampling pattern and the number of solutions of a low-rank tensor completion. In Section III-A, we define a polynomial based on each observed entry. Then, for a given subset of the rank components we transform the problem of finite completability of $\mathcal{U}$ to the problem of finite completability of the core tensor in the Tucker decomposition of $\mathcal{U}$. In Section III-B, we propose a geometric analysis on Tucker manifold, by defining a structure for the core tensor of the Tucker decomposition such that we can determine if two core tensors span the same space. In Section III-C, we construct a constraint tensor based on the sampling pattern Ω. This tensor is useful for analyzing the algebraic independency of a subset of polynomials among all defined polynomials. In Section III-D, we show the relationship between the number of algebraically independent polynomials in the mentioned set of polynomials and finite completability of the sampled tensor. Finally, Section III-E characterizes finite completability in terms of the sampling pattern instead of the algebraic variety for the defined set of polynomials.

A. Condition for Finite Completability Given the Core Tensor

Assume that the sampled tensor is $\mathcal{U} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ and rank components $\{r_{j+1}, \ldots, r_d\}$ are given, where $j \in \{1, 2, \ldots, d-1\}$ is an arbitrary fixed number. Without loss of generality assume that $r_{j+1} \geq \cdots \geq r_d$ throughout the paper. Define $\mathbb{P}_0$ as the Lebesgue measure on $\mathbb{R}^{r_1 \times r_2 \times \cdots \times r_d}$ and $\mathbb{P}_i$ as the Lebesgue measure on $\mathbb{R}^{r_1 \times n_1}$, $i = 1, \ldots, d$. We assume that $\mathcal{U}$ is chosen generically from the manifold corresponding to rank vector $(r_1, \ldots, r_d)$, or in other words, the entries of $\mathcal{U}$ are drawn independently with respect to Lebesgue measure on the corresponding manifold. Hence, any statement that holds for $\mathcal{U}$, it basically holds for almost every (with probability one) tensor of the same size and Tucker-rank with respect to the product measure $\mathbb{P}_0 \times \mathbb{P}_1 \times \cdots \times \mathbb{P}_d$.

Let the dth-order tensor $\mathcal{C} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times r_{j+1} \times r_{j+2} \times \cdots \times r_d}$ be a core tensor of the sampled tensor $\mathcal{U} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$.

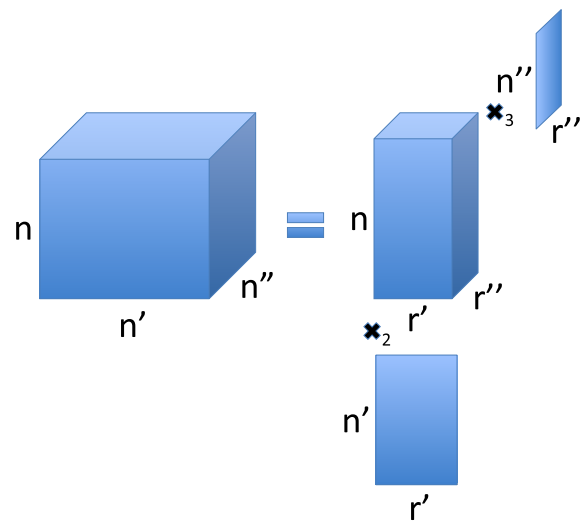

Fig. 1. Tucker decomposition with $j = 1$ and $d = 3$.

Then, there exist full-rank matrices $\mathbf{T}_i$'s with $\mathbf{T}_i \in \mathbb{R}^{r_i \times n_i}$ such that

$$\mathcal{U} = \mathcal{C} \times_{i=j+1}^{d} \mathbf{T}_i, \quad (6)$$

or equivalently

$$\mathcal{U}(\vec{x}) = \sum_{k_{j+1}=1}^{r_{j+1}} \cdots \sum_{k_d=1}^{r_d} \mathcal{C}(x_1, \ldots, x_j, k_{j+1}, \ldots, k_d) \mathbf{T}_{j+1}(k_{j+1}, x_{j+1}) \ldots \mathbf{T}_d(k_d, x_d). \quad (7)$$

For notational simplicity, define $\mathbb{T} = (\mathbf{T}_{j+1}, \ldots, \mathbf{T}_d)$. Figure 1 represents a Tucker decomposition for a 3rd-order tensor given the second and third components of its rank vector.

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

- **Note** 1: As it can be seen from (7), any observed entry $\mathcal{U}(\vec{x})$ results in an equation that involves $\Pi_{i=j+1}^{d} r_i$ entries of $\mathcal{C}$ and also r_i entries of $\mathbf{T}_i$, $i = j+1, \ldots, d$. Considering the entries of core tensor $\mathcal{C}$ and tuple $\mathbb{T}$ as variables (right-hand side of (7)), each observed entry results in a polynomial in terms of these variables.
- **Note** 2: For any observed entry $\mathcal{U}(\vec{x})$, the tuple $(x_1, \ldots, x_j)$ specifies the coordinates of the $\Pi_{i=j+1}^{d} r_i$ entries of $\mathcal{C}$ 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, $i = j+1, \ldots, d$.
- **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 $(\mathcal{C}, \mathbb{T})$ (infinite, finite or unique) via investigating the algebraic independence among these polynomials.
- **Note** 5: Note that it can be concluded from Bernstein's theorem [54] that in a system of n polynomials in n variables with each consisting of a given set of monomials, the n polynomials are algebraically independent with probability one with respect to the corresponding

probability measure, and therefore there exist only finitely many solutions. However, in the structure of the polynomials in our model, the set of involved monomials are different for different set of polynomials, and therefore to ensure algebraically independency we need to have for any selected subset of the original n polynomials, the number of involved variables should be more than the number of selected polynomials.

Given $\mathcal{C}$, we are interested to find a subset of the mentioned polynomials that guarantees tuple $\mathbb{T}$ can be determined finitely. The following definition will be used to determine the number of involved variables in a set of polynomials.

Definition 1: For any $i \in \{j+1, \ldots, d\}$ and nonempty $\mathcal{S}_i \subseteq \{1, \ldots, n_i\}$, define $\mathcal{U}^{(\mathcal{S}_i)}$ as a set containing the locations of the entries of $|\mathcal{S}_i|$ rows (corresponding to the elements of $\mathcal{S}_i$) of $\mathbf{U}_{(i)}$. Moreover, define $\mathcal{U}^{(\mathcal{S}_{j+1},\ldots,\mathcal{S}_d)} = \mathcal{U}^{(\mathcal{S}_{j+1})} \cup \cdots \cup \mathcal{U}^{(\mathcal{S}_d)}$. Let τ be a subset of the locations of the entries of $\mathcal{U}$. Then, $\mathcal{U}^{(\mathcal{S}_{j+1},\ldots,\mathcal{S}_d)}$ 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 locations of the entries in τ.

Example 1. *Consider a tensor $\mathcal{U} \in \mathbb{R}^{4\times4\times4}$ with $j = 1$, $r_2 = 2$ and $r_3 = 3$. Then we have*

$$\mathcal{U}(\vec{x}) = \sum_{k_2=1}^{2}\sum_{k_3=1}^{3} \mathcal{C}(x_1, k_2, k_3)\mathbf{T}_2(k_2, x_2)\mathbf{T}_3(k_3, x_3). \quad (8)$$

where $\mathcal{C} \in \mathbb{R}^{4\times2\times3}$, $\mathbf{T}_2 \in \mathbb{R}^{2\times4}$ and $\mathbf{T}_3 \in \mathbb{R}^{3\times4}$. Define $\tau = \{(1,2,2), (2,2,3), (3,4,2)\}$. Hence, Among the four rows of the second matricization of $\mathcal{U}$, i.e. $\mathbf{U}_{(2)}$, row numbers 2 and 4 include at least one entry in τ (and row numbers 2 and 3 for $\mathbf{U}_{(3)}$). Then, for $\mathcal{S}_2 = \{2,4\}$ $\mathcal{S}_3 = \{2,3\}$, $\mathcal{U}^{(\mathcal{S}_2,\mathcal{S}_3)}$ is the minimal hull of τ. Note that due to the definition, the minimal hull is unique.

Remark 1. Consider any set of polynomials $\{p_1, \ldots, p_k\}$ in form of (7). Given the core tensor $\mathcal{C}$ 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 $\mathcal{U}^{(\mathcal{S}_{j+1},\ldots,\mathcal{S}_d)}$ 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 $\{p_1, \ldots, p_k\}$. 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, $|\mathcal{S}| = \sum_{i=j+1}^{d} |\mathcal{S}_i| r_i$. □

Note that each observed entry results in a scalar equation of the form of (7). Given the core tensor $\mathcal{C}$, 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, 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 nonempty $\mathcal{S}_i \subseteq \{1, \ldots, n_i\}$ for $i \in \{j+1, \ldots, d\}$, $\mathcal{U}^{(\mathcal{S}_{j+1},\ldots,\mathcal{S}_d)}$ includes at most $\sum_{i=j+1}^{d} |\mathcal{S}_i| r_i$ of the mentioned $\sum_{i=j+1}^{d}(n_i r_i)$ observed entries.

Example 2. *Consider a tensor $\mathcal{U} \in \mathbb{R}^{3\times3\times3}$ with $j = 1$, $r_2 = 1$ and $r_3 = 2$. Define the following two sets of observed entries each including $\sum_{i=j+1}^{d}(n_i r_i) = 9$ entries*

$$\begin{aligned}\mathcal{I}_1 = \{&(1,1,1), (1,2,1), (1,1,3), (2,1,1),\\ &(2,2,1), (2,2,3), (2,3,1), (2,3,2), (3,1,1)\},\\ \mathcal{I}_2 = \{&(1,1,1), (1,1,2), (1,2,1), (1,1,3),\\ &(2,2,2), (2,2,3), (2,3,1), (2,3,2), (3,3,3)\}.\end{aligned}$$

First, we can show that $\mathcal{I}_1$ does not satisfy Assumption A_1. For $\mathcal{S}_2 = \{1,2\}$ and $\mathcal{S}_3 = \{1\}$ we have

$$\{(1,1,1), (1,2,1), (2,1,1), (2,2,1), (3,1,1)\} \subset \mathcal{U}^{(\mathcal{S}_2,\mathcal{S}_3)}.$$

Hence, $\mathcal{U}^{(\mathcal{S}_2,\mathcal{S}_3)}$ includes 5 entries belonging to $\mathcal{I}_1$ and $\sum_{i=j+1}^{d} |\mathcal{S}_i| r_i = 4 < 5$. Therefore, $\mathcal{I}_1$ does not satisfy Assumption A_1 since there exists $(\mathcal{S}_2, \mathcal{S}_3)$ that violates the mentioned condition.

Second, it is easy to verify that $\mathcal{I}_2$ satisfies Assumption A_1 by checking all possible pairs $(\mathcal{S}_2, \mathcal{S}_3)$.

Remark 2. *Assume that each column of $\mathbf{U}_{(1)}$ includes at least one observed entry, and also $\sum_{i=j+1}^{d}(n_i r_i) < n_{j+1} \ldots n_d$, for some $j \in \{1, \ldots, d-1\}$. Then, for any tuple $(x_2, x_3, \ldots, x_d)$ that $x_i \in \{1, \ldots, n_i\}$, there exists at least one observed entry among the set*

$$\{(1, x_2, x_3, \ldots, x_d), (2, x_2, x_3, \ldots, x_d), \ldots, (n_1, x_2, x_3, \ldots, x_d)\}. \quad (9)$$

Hence, there exist $\sum_{i=j+1}^{d}(n_i r_i)$ observed entries that satisfy Assumption A_j. This is because all possible tuples $(x_{j+1}, \ldots, x_d)$ are available to be selected. For example, assuming that each column of $\mathbf{U}_{(1)}$ includes at least one observed entry and $\sum_{i=2}^{d}(n_i r_i) < n_2 \ldots n_d$, we can choose $\sum_{i=2}^{d}(n_i r_i) < n_2 \ldots n_d = N_{-1}$ observed entries in different columns of $\mathbf{U}_{(1)}$ to satisfy Assumption A_1, by choosing either zero or one observed entry in each column.

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

Lemma 1. *Assume that in (6) the core tensor $\mathcal{C} \in \mathbb{R}^{n_1\times\cdots\times n_j\times r_{j+1}\times\cdots\times r_d}$ 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 at most finitely many possible tuples $\mathbb{T}$ that satisfy (6) 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 (6)-(7). 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 sufficiency holds.

In order to prove the necessity, 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 lemma. ■

Remark 3. Assumption A_j results that given a core tensor there are finitely many tuples $\mathbb{T}$ such that (6) holds. Consequently, in what follows without loss of generality, we analyze the finite completability of core tensor $\mathcal{C}$ 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 $\mathcal{C}$. Note that since $\mathbb{T}$ is already solved in terms of $\mathcal{C}$, each polynomial in $\mathcal{P}(\Omega)$ is in terms of elements of $\mathcal{C}$.

B. Geometry of Tucker Manifold

We need to define the following equivalence class in order to characterize a condition on the sampling pattern to study the algebraic independency of the polynomials in $\mathcal{P}(\Omega)$. This equivalence class leads to a geometric structure for core tensors which specifies exactly one core tensor among all core tensors that span the same space.

Definition 1. *Define an equivalence class for all core tensors $\mathcal{C} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times r_{j+1} \times \cdots \times r_d}$ of the sampled tensor $\mathcal{U}$ such that two core tensors $\mathcal{C}_1$ and $\mathcal{C}_2$ belong to the same class if and only if there exist full rank matrices $\mathbf{D}_i \in \mathbb{R}^{r_i \times r_i}$, $i = j+1, \ldots, d$, such that*

$$\mathcal{C}_2 = \mathcal{C}_1 \times_{i=j+1}^{d} \mathbf{D}_i. \tag{10}$$

A subtensor $\mathcal{Y} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times 1 \times \cdots \times 1}$ of $\mathcal{U}$ can be represented by a core tensor $\mathcal{C}$ if there exist vectors $\theta_i \in \mathbb{R}^{r_i \times 1}$, $i = j+1, \ldots, d$, such that

$$\mathcal{Y} = \mathcal{C} \times_{i=j+1}^{d} \theta_i. \tag{11}$$

According to Definition 1, it is easy to verify that two core tensors are in the same class if and only if one of them can represent each subtensor in $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times 1 \times \cdots \times 1}$ of the other one.

Definition 2. *Let $N_i = n_1\, n_2 \ldots n_i$, $\bar{N}_i = n_{i+1} n_{i+2} \ldots n_d$ and define the matrix $\widetilde{\mathbf{U}}_{(i)} \in \mathbb{R}^{N_i \times \bar{N}_i}$ as the i-th unfolding of the tensor $\mathcal{U}$, such that $\mathcal{U}(\vec{x}) = \widetilde{\mathbf{U}}_{(i)}(\bar{\mathcal{M}}_{(i)}(x_1, \ldots, x_i), \bar{\bar{\mathcal{M}}}_{(i)}(x_{i+1}, \ldots, x_d))$, where $\bar{\mathcal{M}}_{(i)}$ and $\bar{\bar{\mathcal{M}}}_{(i)}$ are two bijective mappings $\bar{\mathcal{M}}_{(i)} : (x_1, \ldots, x_i) \to \{1, 2, \ldots, N_i\}$ and $\bar{\bar{\mathcal{M}}}_{(i)} : (x_{i+1}, \ldots, x_d) \to \{1, 2, \ldots, \bar{N}_i\}$.*

We make the following assumption which will be referred to, when it is needed.

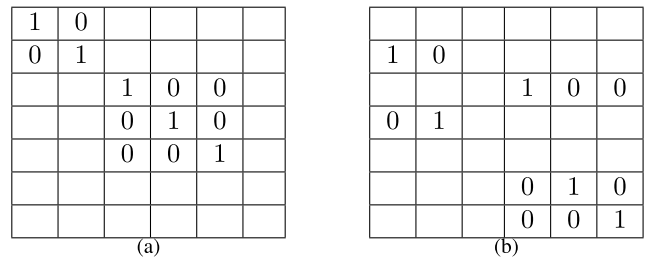

Fig. 2. Two proper structures in $\widetilde{\mathbf{C}}_{(1)}$. (a) Example of canonical core tensor. (b) A proper structure.

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

Consider an arbitrary entry $\mathcal{C}(\vec{x})$ of core tensor $\mathcal{C}$. Note that the tuple $(x_{j+1}, \ldots, x_d)$ specifies the column number of this entry in $\widetilde{\mathbf{C}}_{(j)}$. Furthermore, x_i specifies the row number of this entry in $\mathbf{C}_{(i)}$. Consequently, each column of $\widetilde{\mathbf{C}}_{(j)}$ indexed by $(x_{j+1}, \ldots, x_d)$ belongs to the x_i^{th} row of $\mathbf{C}_{(i)}$ for $i \geq j+1$.

We are interested in providing a structure on the core tensor $\mathcal{C}$ such that exactly one core tensor in any class satisfies it. We present this structure on $\widetilde{\mathbf{C}}_{(j)}$ (j-th unfolding of $\mathcal{C}$).

Definition 3. *Consider any $(d-j)$ disjoint submatrices*[1] *$(\mathbf{P}_{j+1}, \ldots, \mathbf{P}_d)$ of $\widetilde{\mathbf{C}}_{(j)}$ such that (i) $\mathbf{P}_i \in \mathbb{R}^{r_i \times r_i}$, $i = j+1, \ldots, d$, (ii) the $\sum_{i=j+1}^{d} r_i$ rows of $\widetilde{\mathbf{C}}_{(j)}$ corresponding to rows of these submatrices are disjoint, (iii) the r_i columns of $\widetilde{\mathbf{C}}_{(j)}$ corresponding to columns of $\mathbf{P}_i$ belong to r_i disjoint rows of $\mathbf{C}_{(i)}$, $i = j+1, \ldots, d$. Then, $\mathcal{C}$ is said to have a proper structure if $\mathbf{P}_i = \mathbf{I}_{r_i}$, $i = j+1, \ldots, d$.*

Assumption B_j ensures the existence of a proper structure since the number of rows in $\widetilde{\mathbf{C}}_{(j)}$ should be at least $\sum_{i=j+1}^{d} r_i$. Note that given a proper structure there exists one core tensor in each class that satisfies it. Moreover, we can permute the rows of $\mathbf{P}_i$'s and obtain another proper structure. Consider the following specific structure. Define

$$\mathbf{P}_i^{\text{can}}(x_1', x_i) = \mathcal{C}(x_1, \underbrace{1, \ldots, 1}_{i-2}, x_i, \underbrace{1, \ldots, 1}_{d-i}) \in \mathbb{R}^{r_i \times r_i}, \quad i = j+1, \ldots, d, \tag{12}$$

where $x_1' = x_1 - \sum_{s=j+1}^{i-1} r_s$, $1 + \sum_{s=j+1}^{i-1} r_s \leq x_1 \leq \sum_{s=j+1}^{i} r_s$ and $1 \leq x_i \leq r_i$. It is easily verified all three properties in Definition 3 are satisfied by $\mathbf{P}_i^{\text{can}}$, $i = j+1, \ldots, d$.

Definition 4. (Canonical core tensor) *We call $\mathcal{C}$ a canonical core tensor if for $i \geq j+1$ we have $\mathbf{P}_i^{can} = \mathbf{I}_{r_i}$, where $\mathbf{I}_{r_i}$ is the $r_i \times r_i$ identity matrix.*

Example 3. Assume $\mathcal{U} \in \mathbb{R}^{7 \times 3 \times 4}$, $j = 1$, $r_2 = 2$ and $r_3 = 3$. For simplicity in representing the canonical core tensor, we consider bijective mappings $\bar{\mathcal{M}}_{(i)}$ and $\bar{\bar{\mathcal{M}}}_{(i)}$ (in Definition 2) that result in the structure of $\widetilde{\mathbf{C}}_{(1)}$ shown in Figure 2(a). Observe that any permutation of rows of $\widetilde{\mathbf{C}}_{(1)}$ results in a proper structure, e.g., Figure 2(b). However, only those permutations of columns of $\widetilde{\mathbf{C}}_{(1)}$ that satisfy property (iii) in Definition 3 result in a proper structure. □

[1] Specified by a subset of rows and a subset of columns (not necessarily consecutive).

Note that (10) leads to the fact that the dimension of all core tensors $\mathcal{C}$ that span different spaces (without any polynomial restrictions in $\mathcal{P}(\Omega)$) is equal to $\left(\Pi_{i=1}^{j} n_i\right)\left(\Pi_{i=j+1}^{d} r_i\right) - \left(\sum_{i=j+1}^{d} r_i^2\right)$, as the total number of entries of $\mathbf{D}_i$'s is equal to $\left(\sum_{i=j+1}^{d} r_i^2\right)$. Moreover, observe that a core tensor with a proper structure has $\left(\sum_{i=j+1}^{d} r_i^2\right)$ known entries, and therefore the number of unknown entries is equal to $\left(\Pi_{i=1}^{j} n_i\right)\left(\Pi_{i=j+1}^{d} r_i\right) - \left(\sum_{i=j+1}^{d} r_i^2\right)$.

Remark 4. In order to prove there are finitely many completions for tensor $\mathcal{U}$, it suffices to prove that there are finitely many canonical core tensors that fit in $\mathcal{U}$. □

Suppose $\mathcal{C}$ has a proper structure. Let $g_{j+1}(x)$ denote the maximum number of known entries among any x rows of $\widetilde{\mathbf{C}}_{(j)}$. As will be seen in Section III-C, $g_{j+1}(x)$ plays an important role in expressing the maximum number of algebraically independent polynomials in a subset of $\mathcal{P}(\Omega)$. Note that in exactly r_i rows of $\widetilde{\mathbf{C}}_{(j)}$ there are exactly r_i known entries, i.e., entries of $\mathbf{P}_i$, $i = j+1, \ldots, d$. Also, there are $\sum_{i=j+1}^{d} r_i$ rows that include known entries in $\widetilde{\mathbf{C}}_{(j)}$, i.e., rows of $\mathbf{P}_i$, $i = j+1, \ldots, d$.

Recall the assumption $r_{j+1} \geq \cdots \geq r_d$. Therefore, as long as $x \leq r_{j+1}$, the maximum number of known entries is $g_{j+1}(x) = r_{j+1}x$ by selecting the x rows of $\widetilde{\mathbf{C}}_{(j)}$ to cover any x rows of $\mathbf{P}_{j+1}$. On the other hand, if $r_{j+1} \leq x \leq r_{j+1} + r_{j+2}$, the maximum number of known entries is $g_{j+1}(x) = r_{j+1}^2 + r_{j+2}(x - r_{j+1})$ by selecting the x rows of $\widetilde{\mathbf{C}}_{(j)}$ to cover all rows of $\mathbf{P}_{j+1}$ and any $(x - r_{j+1})$ rows of $\mathbf{P}_{j+2}$. Then, in general we have

$$g_{j+1}(x) = \sum_{i=j+1}^{d} \min\left\{r_i, \left(x - \sum_{i'=j+1}^{i-1} r_{i'}\right)^{+}\right\} r_i. \quad (13)$$

C. Constraint Tensor

In the following, we propose a procedure to construct a $(j+1)^{\text{th}}$-order binary tensor $\breve{\mathbf{\Omega}}$ based on Ω such that $\mathcal{P}(\breve{\mathbf{\Omega}}) = \mathcal{P}(\mathbf{\Omega})$. Using $\breve{\mathbf{\Omega}}$, we are able to recognize the observed entries that have been used to obtain the tuple $\mathbb{T}$, and we can easily verify if two polynomials in $\mathcal{P}(\mathbf{\Omega})$ are in terms of the same set of variables. Then, in Section III-D, we characterize the relationship between the maximum number of algebraically independent polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}})$ and $\breve{\mathbf{\Omega}}$.

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 (11) holds or equivalently

$$\begin{aligned}\mathcal{Y}(x_1, \ldots, x_j, \vec{1}_{d-j}) &= \sum_{k_{j+1}=1}^{r_{j+1}} \cdots \sum_{k_d=1}^{r_d} \mathcal{C}(x_1, \ldots, x_j, k_{j+1}, \ldots, k_d) \\ &\quad \times \theta_{j+1}(k_{j+1}, 1) \ldots \theta_d(k_d, 1), \end{aligned} \quad (14)$$

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 core tensor $\mathcal{C}$ among all $N_\Omega(\mathcal{U}) - \sum_{i=j+1}^{d}(n_i r_i)$ polynomials in $\mathcal{P}(\mathbf{\Omega})$.

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 $\breve{\mathcal{Y}}_{(t_{j+1}, \cdots, t_d)} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times \overbrace{1 \times \cdots \times 1}^{d-j} \times k}$, 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 (6) 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 order to shed some light on the above procedure we give an illustrative example in the following.

Example 4. Consider an example in which $d = 3$, $j = 2$, $r_3 = 2$ and $\mathcal{U} \in \mathbb{R}^{3 \times 2 \times 2}$. Assume that $\Omega(x, y, z) = 1$ if $(x, y, z) \in \mathcal{S}$ and $\Omega(x, y, z) = 0$ otherwise, where

$$\mathcal{S} = \{(1,1,1), (1,2,1), (2,2,1), (3,1,1), (1,1,2), (2,1,2), (3,2,2)\},$$

represents the set of observed entries. Also, assume that $N_\Omega(\mathcal{Y}^{\mathbb{T}}_{(1)}) = N_\Omega(\mathcal{Y}^{\mathbb{T}}_{(2)}) = 2$ and the entries that we use to obtain $\mathbb{T}$ are $(1,1,1)$, $(1,2,1)$, $(1,1,2)$ and $(3,2,2)$. Hence, $\breve{\mathcal{Y}}_{(1)} \in \mathbb{R}^{3 \times 2 \times 1 \times 2}$, $\breve{\mathcal{Y}}_{(2)} \in \mathbb{R}^{3 \times 2 \times 1 \times 1}$, and therefore the constraint tensor $\breve{\mathbf{\Omega}}$ belongs to $\mathbb{R}^{3 \times 2 \times 3}$.

Note that $\breve{\mathcal{Y}}_{(1)}(1,1,1,1) = \breve{\mathcal{Y}}_{(1)}(1,1,1,2) = \breve{\mathcal{Y}}_{(1)}(1,2,1,1) = \breve{\mathcal{Y}}_{(1)}(1,2,1,2) = 1$, and also for the two other observed entires we have $\breve{\mathcal{Y}}_{(1)}(2,2,1,1) = 1$ and $\breve{\mathcal{Y}}_{(1)}(3,1,1,2) = 1$ and the rest of the entries of $\breve{\mathcal{Y}}_{(1)}$ are equal to zero. Moreover, $\breve{\mathcal{Y}}_{(2)}(1,1,1,1) = \breve{\mathcal{Y}}_{(2)}(2,1,1,1) = \breve{\mathcal{Y}}_{(2)}(3,2,1,1) = 1$ and the rest of the entries of $\breve{\mathcal{Y}}_{(2)}$ are equal to zero.

Then, $\breve{\mathbf{\Omega}}(x, y, z) = 1$ if $(x, y, z) \in \mathcal{S}'$ and $\breve{\mathbf{\Omega}}(x, y, z) = 0$ otherwise, where

$$\breve{\mathcal{S}} = \{(1,1,1), (1,2,1), (2,2,1), (1,1,2), (1,2,2), (3,1,2), (1,1,3), (3,2,3), (2,1,3)\}.$$

□

Note that each subtensor of $\breve{\mathbf{\Omega}}$ that belongs to $\mathbb{R}^{n_1 \times \cdots \times n_j \times 1}$ represents one of the polynomials in $\mathcal{P}(\mathbf{\Omega})$ besides showing the

polynomials that have been used to obtain $\mathbb{T}$. More specifically, consider a subtensor of $\breve{\mathbf{\Omega}}$ that belongs to $\mathbb{R}^{n_1\times\cdots\times n_j\times 1}$ with $l+1$ nonzero entries. Observe that exactly l of them correspond to the observed entries that have been used to obtain $\mathbb{T}$. Hence, this subtensor represents a polynomial after replacing entries of $\mathbb{T}$ by the expressions in terms of entries of $\mathcal{C}$, i.e., a polynomial in $\mathcal{P}(\Omega)$.

Recall that the tuple $(x_{j+1},\ldots,x_d)$ specifies the column number of entry $\mathcal{C}(\vec{x})$ in $\widetilde{\mathbf{C}}_{(j)}$. Hence, if the i-th column of $\breve{\mathbf{\Omega}}_{(j+1)}$ is nonzero, then there exists a polynomial in $\mathcal{P}(\Omega)$ that involves all entries of core tensor corresponding to the entries of the i-th row of $\widetilde{\mathbf{C}}_{(j)}$.

D. 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 $\mathcal{P}(\breve{\mathbf{\Omega}})$.

The following lemma determines the required number of algebraically independent polynomials in $\mathcal{P}(\Omega)$ that is needed to ensure finite completability of the core tensor.

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: Assume that $\mathcal{C}\in\mathbb{R}^{n_1\times n_2\times\cdots\times n_j\times r_{j+1}\times\cdots\times r_d}$ is a core tensor for the sampled tensor $\mathcal{U}$. Since assumption A_j holds, Lemma 1 results that there exist finitely many tuples $\mathbb{T}$ such that (6) holds. However, according to Remark 3, it suffices to assume $\mathbb{T}$ is fixed and then prove the statement. Let $\mathcal{P}(\Omega)=\{p_1,\ldots,p_m\}$ and define $\mathcal{S}_i$ as the set of all core tensors that satisfy polynomial restrictions $\{p_1,\ldots,p_i\}$, $i=0,\ldots,m$ ($\mathcal{S}_0$ is the set of all core tensors without any polynomial restriction).

Observe that each algebraically independent polynomial reduces the dimension (degree of freedom) of the set of solutions by one. In other words, $\dim(\mathcal{S}_i)=\dim(\mathcal{S}_{i-1})$ if the maximum number of algebraically independent polynomials in sets $\{p_1,\ldots,p_i\}$ and $\{p_1,\ldots,p_{i-1}\}$ are the same and $\dim(\mathcal{S}_i)=\dim(\mathcal{S}_{i-1})-1$ otherwise. Moreover, with probability one, $|\mathcal{S}_m|$ is finite if and only if there are $\dim(\mathcal{C})=\dim(\mathcal{S}_0)$ algebraically independent polynomial restrictions on the entries of the core tensor $\mathcal{C}$, i.e., $|\mathcal{S}_m|$ is finite if and only if $\dim(\mathcal{S}_m)=0$ [34]. Hence, there are finitely many completions of the sampled tensor $\mathcal{U}$ if and only if there exist $\dim(\mathcal{C})=\left(\Pi_{i=1}^{j} n_i\right)\left(\Pi_{i=j+1}^{d} r_i\right)-\left(\sum_{i=j+1}^{d} r_i^2\right)$ algebraically independent polynomials in $\mathcal{P}(\Omega)$. ■

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 5. *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 $\breve{\mathbf{\Omega}}'$.*

Recall Note 2 regarding the number of involved entries of core tensor in a set of polynomials. However, as mentioned earlier, some of the entries of $\mathcal{C}$ are known, i.e., $(\mathbf{P}_{j+1},\ldots,\mathbf{P}_d)$. Therefore, in order to find the number of variables (unknown entries of $\mathcal{C}$) in a set of polynomials, we should subtract the number of known entries in the corresponding pattern from the total number of involved entries.

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.

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*

$$\left(\Pi_{i=j+1}^{d} r_i\right) m_{j+1}(\breve{\mathbf{\Omega}}')-g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}')), \tag{15}$$

where $g_{j+1}(\cdot)$ is given in (13).

Proof: Observe that the number of algebraically independent polynomials in a subset of polynomials of $\mathcal{P}(\breve{\mathbf{\Omega}}')$ is at most equal to the total number of variables that are involved in the corresponding polynomials. According to (7), for each observed entry $\mathcal{U}(\vec{x})$, we have a polynomial in terms of $\Pi_{i=j+1}^{d} r_i$ entries of the core tensor $\mathcal{C}$ corresponding to the first j coordinates of the location of the observed entry. Therefore, the number of such variables in $\mathcal{P}(\breve{\mathbf{\Omega}}')$ will be $\Pi_{i=j+1}^{d} r_i$ times the number of different tuples $(x_1,\ldots,x_j)$ among the corresponding observed entries. The number of nonzero columns of $\breve{\mathbf{\Omega}}'_{(j+1)}$ is exactly equal to the number such tuples. Therefore, the number of involved entries of $\mathcal{C}$ (known and unknown) in polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}}')$ is equal to $\left(\Pi_{i=j+1}^{d} r_i\right) m_{j+1}(\breve{\mathbf{\Omega}}')$

On the other hand, among the $\sum_{i=j+1}^{d} r_i^2$ known entries corresponding to $(\mathbf{P}_{j+1},\ldots,\mathbf{P}_d)$ in $\widetilde{\mathbf{C}}_{(j)}$, $g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}'))$ of them are involved in polynomials of $\mathcal{P}(\breve{\mathbf{\Omega}}')$. This is because entries of the i-th row of $\widetilde{\mathbf{C}}_{(j)}$ are involved in a polynomials if and only if the i-th column of $\breve{\mathbf{\Omega}}'_{(j+1)}$ includes at least one nonzero entry. Hence, the number of variables that are involved in the set of polynomials $\mathcal{P}(\breve{\mathbf{\Omega}}')$ is given by (15) for a particular proper structure and proof is complete as (15) is an upper bound for the number of algebraically independent polynomials. ■

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. The following lemma is a re-statement of Lemma 7 in [56].

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 $t-1$.*

Finally, the next theorem provides a relationship between the exact number of algebraically independent polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}})$ and a geometric property on $\breve{\mathbf{\Omega}}$.

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 $\mathcal{P}(\breve{\mathbf{\Omega}}')$ since $g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}'))$ is the maximum number of known entries of core tensor 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. ■

E. 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 and only 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

$$\left(\Pi_{i=j+1}^{d} r_i\right) m_{j+1}(\breve{\mathbf{\Omega}}'') - g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}'')) \geq t. \tag{16}$$

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. ■

Remark 5. *As a sanity check, we next show that when $d = 2$ (matrix case), Theorem 3 reduces to Theorem 1 in [34]. To see this, note that for $d = 2$, we only have one rank component and denote it by r. Then, Condition (i) states that there exists a submatrix $\breve{\Omega}' \in \mathbb{R}^{n_1 \times (n_1 r - r^2)}$ (basically $n = n_1 r - r^2$ in Condition (i)) of the constraint matrix $\breve{\Omega}$. And Condition (ii) on the property of submatrix $\breve{\Omega}'$ becomes: for any $t \in \{1, \ldots, n_1 r - r^2\}$ and any submatrix $\breve{\mathbf{\Omega}}'' \in \mathbb{R}^{n_1 \times t}$ of the matrix $\breve{\mathbf{\Omega}}'$, the following inequality holds*

$$r m_1(\breve{\mathbf{\Omega}}'') - g_1(m_1(\breve{\mathbf{\Omega}}'')) \geq t. \tag{17}$$

Note that due to the way that we constructed the constraint matrix (tensor) each column of $\breve{\mathbf{\Omega}}''$ has exactly $r+1$ non-zero entries. Therefore, $\breve{\mathbf{\Omega}}''$ has at least $r+1$ non-zero rows, i.e., $m_1(\breve{\mathbf{\Omega}}'') \geq r+1$. Then, according to the definition of g_1 in (13)*, $g_1(m_1(\breve{\mathbf{\Omega}}'')) = r \min\{r, m_1(\breve{\mathbf{\Omega}}'') - 0\} = r^2$. Therefore,* (17) *becomes*

$$r m_1(\breve{\mathbf{\Omega}}'') - r^2 \geq t, \tag{18}$$

or equivalently,

$$m_1(\breve{\mathbf{\Omega}}'') \geq t/r + r. \tag{19}$$

Hence, Theorem 3 for $d = 2$ is exactly the same as Theorem 1 in [34].

There are a few observations based on Theorem 3:

- In theorem 3, we characterize all of the sampling patterns that ensure there are only finitely many completions such that if only one single sample from that pattern is missed, then there are infinitely many completions.
- If we use the Grassmannian analysis on each dimension individually, we will only obtain a *sufficient* condition on the sampling patterns for finite completability, given multiple rank constraints. However, our proposed analysis on Tucker manifold results in a *necessary and sufficient* condition on the sampling patterns for finite completability when multiple rank components are given. Hence, our proposed analysis on Tucker manifold in general requires much less number of samples to ensure finite completability in comparison with the analysis on Grassmannian manifold when more than one rank component is given.
- It is also important to observe the advantage of the proposed method as we decrease the value of j since

we are incorporating more rank components. Intuitively, for the case of $j = 1$, the polynomials obtained in (6) involve much more variables in comparison with the low-order analysis on Grassmannian manifold, i.e., $j = d - 1$. Therefore, the case of $j = 1$ requires much less number of samples in order to have sufficient number of algebraically independent polynomials for finite completability.

- Note that our analysis is not valid when $j = 0$ (given all rank components) since the defined proper structure does not have meaning any more as zero unfolding does not exist, and therefore we are not able to characterize the necessary and sufficient condition on sampling pattern for finite completability. However, if all rank components are given we can simply ignore one of them and characterize the necessary and sufficient condition for $j = 1$ which results in a sufficient condition for $j = 0$.

If the number of completions given $(r_{j+1}, \ldots, r_d)$ is finite, it can be concluded that the number of completions given $(r_j, \ldots, r_d)$ is finite as well. Therefore, this property should be verifiable through the geometric property (16) proposed in Theorem 3. In the following lemma, using the pigeonhole principle, we show this result without analyzing the algebraic variety.

Lemma 4. *Assume that Assumptions A_j, B_j, A_{j-1} and B_{j-1} hold. Suppose that the sampling pattern is such that the properties (i) and (ii) in the statement of Theorem 3 hold for j. Then, the properties (i) and (ii) in the statement of Theorem 3 hold for $j - 1$.*

Proof: The proof is given in Appendix A. ■

IV. Probabilistic Conditions for Finite Completability

In this section, two different lower bounds on the sampling probability are proposed and analyzed to ensure finite completability. The first bound is obtained by applying [34, Theorem 3] and the second bound is obtained through the proposed geometric approach in Theorem 3. We will observe later that our proposed analysis on Tucker manifold leads to a better lower bound through numerical analysis.

Here we briefly outline the key steps of the second approach. Lemmas 6 and 7 each provides a lower bound on the sampling probability that results in a geometric property for Ω, i.e., inequalities (23) and (30), respectively. Then, Theorem 5 takes advantage of the above lemmas to propose a bound on the sampling probability to guarantee property (16) for Ω. Finally, Lemma 8 shows that (16) also holds for the constraint tensor $\breve{\mathbf{\Omega}}$. In order to show Lemma 8 we develop a generalization of Hall's Theorem on bipartite graphs in Theorem 6.

We use the approach similar to [34, Lemma 9] in order to apply Theorem 3 and obtain a lower bound on the sampling probability to ensure that there are only finitely many completions. According to our earlier discussion, Theorem 3 for the case of $j = 1$ results in the mildest condition on the sampling patterns for ensuring finite completability. In other words, setting $j = 1$ in Theorem 3 results in a tighter lower bound.

A. Lower Bound on Sampling Probability Based on Analysis on Grassmannian Manifold

We first state a lemma, whose corollary (Corollary 1) is used extensively in this section in order to find a lower bound on the sampling probability to guarantee a lower bound on the number of sampled entries with high probability.

Lemma 5. *Consider a vector with n_i entries and assume each entry is observed with probability p and independently from the other entries. Then, with probability at least $1-\exp(-\frac{n_i}{2c^2})$ at least $(p - \frac{1}{c})n_i$ entries are observed.*

Proof: Azuma's inequality states that for a martingale $\{X_k : k = 0, 1, 2, ...\}$ that $|X_k - X_{k-1}| \leq 1$ holds, we have $P(X_n - \mathbb{E}[X_n] > t) \leq \exp\left(\frac{-t^2}{2n}\right)$ and $P(\mathbb{E}[X_n] - X_n > t) \leq \exp\left(\frac{-t^2}{2n}\right)$ [57]. Therefore, using Azuma's inequality and the fact that sampling has Bernoulli distribution with parameter p, it can be seen that with probability at most $\exp(-\frac{t^2}{2n_i})$ the number of observed entries is less than $n_i p - t$. Now, by setting t equal to $\frac{n_i}{c}$ the proof is complete. ■

Corollary 1. *Consider a vector with n_i entries where each entry is observed with probability p independently from the other entries. If $p > p' = \frac{2r_i}{n_i} + \frac{1}{\sqrt[4]{n_i}}$, with probability at least $\left(1 - \exp(-\frac{\sqrt{n_i}}{2})\right)$, more than $2r_i$ entries are observed. Similarly, if $p > p'' = \frac{12\log(\frac{en_i}{\epsilon})}{n_i} + \frac{1}{\sqrt[4]{n_i}}$, with probability at least $\left(1 - \exp(-\frac{\sqrt{n_i}}{2})\right)$, more than $12\log(\frac{en_i}{\epsilon})$ entries are observed.*

Proof: Both parts follow by setting $c = \sqrt[4]{n_i}$ in Lemma 5. ■

The following theorem uses the matrix result for single rank constraint and provides a sufficient condition on the sampling pattern for finite completabilitily of tensor. We will assume that

$$r_i \leq \frac{n_i}{6}, \quad N_{-i} \geq r_i(n_i - r_i) \quad \text{for} \quad i = 1, 2, \ldots, d. \tag{20}$$

Theorem 4. *Consider a randomly sampled tensor $\mathcal{U}$ with Tucker rank $(r_1, \cdots, r_d)$. Suppose that the inequalities in (20) hold. Moreover, assume that the sampling probability satisfies*

$$p > \min_{1 \leq i \leq d} \left(\max\left(\frac{2r_i}{n_i}, \frac{12\log(\frac{en_i}{\epsilon})}{n_i}\right) + \frac{1}{\sqrt[4]{n_i}}\right). \tag{21}$$

Then, there are only finitely many completions of the sampled tensor $\mathcal{U}$ with the given rank vector with probability at least $(1 - 2\epsilon)$.

Proof: Denote $\ell = \arg\min_{1 \leq i \leq d} \left(\max\left(\frac{2r_i}{n_i}, \frac{12\log(\frac{en_i}{\epsilon})}{n_i}\right) + \frac{1}{\sqrt[4]{n_i}}\right)$. By (21) and using Corollary 1, with probability almost one the number of observed entries at each column of $\mathbf{U}_{(\ell)}$ is at least $\max\{2r_\ell, 12\log(\frac{en_\ell}{\epsilon})\}$. Then, with (20) and using [34, Theorem 3], it follows that there are finitely many matrix completions for the observed entries of $\mathbf{U}_{(\ell)}$ with

rank r_ℓ. As a result, there are finitely many tensors with the ℓ-th component of rank being r_ℓ that agree with the observed entries. The proof is complete since the set of tensors that are a completion with the rank vector $(r_1, \ldots, r_d)$ is a subset of the set of tensors whose finiteness is shown above. ■

B. Lower Bound on Sampling Probability Based on Analysis on Tucker Manifold

We are interested in taking advantage of Theorem 3 to obtain another lower bound on the sampling probability for finite completability. We assume $j = 1$, and therefore the constraint tensor $\check{\mathbf{\Omega}}$ is a second-order tensor, i.e., an $n_1 \times K_1$ matrix.

Lemma 6. *Consider an arbitrary set $\Omega'_{(1)}$ of $n_1 - 1$ columns of $\Omega_{(1)}$ (first matricization of Ω). Assume that each column of $\Omega_{(1)}$ includes at least l nonzero entries, where*

$$l > 6 \ \log(n_1) + 2 \ \log\left(\frac{k}{\epsilon}\right) + 4. \tag{22}$$

Then, with probability at least $1 - \frac{\epsilon}{k}$, every subset $\Omega''_{(1)}$ of columns of $\Omega'_{(1)}$ satisfies

$$m_2(\Omega''_{(1)}) - 1 \geq t, \tag{23}$$

where t is the number of columns of $\Omega''_{(1)}$ and $m_2(\Omega''_{(1)})$ is the number of nonzero rows of $\Omega''_{(1)}$ (observe that second matricization of a matrix is its transpose).

Proof: The proof is similar to the proof of [34, Lemma 9] with some delicate modifications to improve the result for this case. Note that (22) can be rewritten as

$$l > 2 \ \log\left(\frac{n_1 k}{\epsilon}\right) + 4 \ \log(n_1) + 4 \tag{24}$$

$$> \log\left(\frac{n_1 e^2 k}{\epsilon}\right) + 2. \tag{25}$$

Define $\mathcal{E}$ as the event that for some submatrix $\Omega''_{(1)} \in \mathbb{R}^{n_1 \times t}$ of the matrix $\Omega'_{(1)}$ (23) does not hold. We are interested in finding an upper bound on the probability of $\mathcal{E}$. Then, from the proof of [34, Lemma 9] and by setting $r = 1$ in inequalities (12) and (13) in [34], we have

$$P(\mathcal{E}) < \sum_{n=l}^{\frac{n_1}{2}} \binom{n_1}{n}^2 \left(\frac{n}{n_1}\right)^{l n_1} + \sum_{n=1}^{\frac{n_1}{2}} \binom{n_1}{n_1 - n}^2 \left(\frac{n_1 - n}{n_1}\right)^{l(n_1 - n)}. \tag{26}$$

Note that

$$\binom{s}{q} = \frac{s(s-1)\ldots(s-q+1)}{q!} \leq \frac{s^q}{q!} \leq \left(\frac{se}{q}\right)^q, \tag{27}$$

where the last inequality holds since $e^q = \sum_{i=0}^{\infty} \frac{q^i}{i!} \leq \frac{q^q}{q!}$. Applying (27) to each term of the first summation of (26) we obtain

$$\binom{n_1}{n}^2 \left(\frac{n}{n_1}\right)^{l n_1} < e^{2n} \left(\frac{n}{n_1}\right)^{(l-2)n} \leq \left(e^2 2^{-l+2}\right)^n < \frac{\epsilon}{n_1 k}, \tag{28}$$

where the last inequality follows from (25). This directly results that the first summation in (26) is less than $\frac{\epsilon}{2k}$.

Moreover, again by applying (27) to each term of the second summation in (26) we obtain

$$\binom{n_1}{n_1 - n}^2 \left(\frac{n_1 - n}{n_1}\right)^{l(n_1 - n)} = \binom{n_1}{n}^2 \left(\frac{n_1 - n}{n_1}\right)^{l(n_1 - n)} < (n_1 e)^{2n} \left(\frac{n_1 - n}{n_1}\right)^{\frac{l n_1}{2}} \leq (n_1 e)^{2n} e^{-\frac{l n}{2}} = \left(e^{2\log n_1 + 2 - \frac{l}{2}}\right)^n < \frac{\epsilon}{n_1 k}, \tag{29}$$

where the last inequality follows from (24). This directly results that the second summation in (26) is less than $\frac{\epsilon}{2k}$ and that completes the proof. ■

Lemma 7. *Consider an arbitrary set $\Omega'_{(1)}$ of n_1 columns of $\Omega_{(1)}$ (first matricization of Ω). Assume that each column of $\Omega_{(1)}$ includes at least l nonzero entries, where $l > 6 \ \log(n_1) + 2 \ \log\left(\frac{2k}{\epsilon}\right) + 4$. Then, with probability at least $1 - \frac{\epsilon}{2k}$, every subset $\Omega''_{(1)}$ of columns of $\Omega'_{(1)}$ satisfies*

$$m_2(\Omega''_{(1)}) \geq t, \tag{30}$$

where t is the number of columns of $\Omega''_{(1)}$ and $m_2(\Omega''_{(1)})$ is the number of nonzero rows of $\Omega''_{(1)}$.

Proof: By setting $r = 0$ in inequalities (12) and (13) in [34], we have

$$P(\mathcal{E}) < \sum_{n=l}^{\frac{n_1}{2}} \binom{n_1}{n}^2 \left(\frac{n}{n_1}\right)^{l(n_1+1)} + \sum_{n=1}^{\frac{n_1}{2}} \binom{n_1}{n_1 - n}^2 \left(\frac{n_1 - n}{n_1}\right)^{l(n_1 - n + 1)}. \tag{31}$$

Since $n < n_1$ it then follows that $P(\mathcal{E}) <$ RHS of (31) $<$ RHS of (26) $< \frac{\epsilon}{2k}$, where the last inequality follows from Lemma 6. ■

The next theorem gives a lower bound on the number of needed samples to ensure finite completability for tensor $\mathcal{U}$. More specifically, we show under some mild assumptions if (32) holds, all conditions and assumptions in the statement of Theorem 3 hold with high probability, and therefore the tensor is finitely completable.

Theorem 5. *Assume that $\sum_{i=2}^d (n_i r_i) < n_2 \ldots n_d$, $\sum_{i=2}^d r_i^2 \leq \Pi_{i=2}^d r_i$, $\Pi_{i=2}^d n_i \geq n_1 \Pi_{i=2}^d r_i - \sum_{i=2}^d r_i^2$, and Assumption B_1 hold. Furthermore, assume that each column of $\mathbf{U}_{(1)}$ includes at least l observed entries, where*

$$l > 6 \ \log(n_1) + 2 \max\left\{\log\left(\frac{2\sum_{i=2}^d r_i^2}{\epsilon}\right), \log\left(\frac{2\Pi_{i=2}^d r_i - 2\sum_{i=2}^d r_i^2}{\epsilon}\right)\right\} + 4. \tag{32}$$

Then, with probability at least $1 - \epsilon$, there exist only finitely many completions, given the rank components $r_2, \ldots, r_d$.

Proof: Since each column of $\Omega_{(1)}$ includes at least one nonzero entry, there exist $\sum_{i=2}^d n_i r_i$ entries in different columns such that they satisfy Assumption A_1 (as mentioned

in Remark 2). Also, $\Omega_{(1)}$ has $\Pi_{i=2}^d n_i$ columns and by assumption $\Pi_{i=2}^d n_i \geq n_1 \Pi_{i=2}^d r_i - \sum_{i=2}^d r_i^2$, therefore there exist $\Pi_{i=2}^d r_i$ disjoint sets $\mathcal{S}_1, \ldots, \mathcal{S}_{\Pi_{i=2}^d r_i}$ such that each $\mathcal{S}_\ell$ consists of $n_1 - 1$ columns of $\Omega_{(1)}$ for $1 \leq \ell \leq \sum_{i=2}^d r_i^2$ and each $\mathcal{S}_\ell$ consists of n_1 columns of $\Omega_{(1)}$ for $\sum_{i=2}^d r_i^2 + 1 \leq \ell \leq \Pi_{i=2}^d r_i$.

Note that by (32), we have $l > 6 \log(n_1) + 2\log\left(\frac{2\sum_{i=2}^d r_i^2}{\epsilon}\right) + 4$. Using Lemma 6, with $k = 2\sum_{i=2}^d r_i^2$, for each $1 \leq \ell \leq \sum_{i=2}^d r_i^2$, with probability at least $1 - \frac{\epsilon}{2\sum_{i=2}^d r_i^2}$, (23) holds for every subset of columns of $\mathcal{S}_\ell$. Hence, with probability at least $1 - \frac{\epsilon}{2}$, (23) holds for every subset of columns of $\mathcal{S}_\ell$ for $1 \leq \ell \leq \sum_{i=2}^d r_i^2$, simultaneously. According to Remark 8 and Lemma 8 below and by setting $r = 1$, as a subset of columns $\mathcal{S}_\ell$ of $\Omega_{(1)}$ satisfies (23), there exist a subset of columns $\breve{\mathcal{S}}_\ell$ of the constraint matrix $\breve{\mathbf{\Omega}}$ (corresponding columns to the columns of $\mathcal{S}_\ell$) that satisfies (23) as well, for $1 \leq \ell \leq \sum_{i=2}^d r_i^2$. Denote $\breve{\mathcal{S}}^{(0)} = \cup_{\ell=1}^{\sum_{i=2}^d r_i^2} \breve{\mathcal{S}}_\ell$.

Consider any subset of columns of $\breve{\mathcal{S}}^{(0)}$ and denote it by $\breve{\mathcal{S}}'$. Let $\breve{\mathcal{S}}'_\ell$ denote the columns of $\breve{\mathcal{S}}'$ that belong in $\breve{\mathcal{S}}_\ell$ and, without loss of generality, assume that $|\breve{\mathcal{S}}'_1| \geq \cdots \geq |\breve{\mathcal{S}}'_{\sum_{i=2}^d r_i^2}|$, where $|\breve{\mathcal{S}}'_\ell|$ denotes the number of columns of $\breve{\mathcal{S}}'_\ell$. As a result, we obtain

$$|\breve{\mathcal{S}}'| = \sum_{\ell=1}^{\sum_{i=2}^d r_i^2} |\breve{\mathcal{S}}'_\ell| \leq \left(\sum_{i=2}^d r_i^2\right) |\breve{\mathcal{S}}'_1| \leq \left(\sum_{i=2}^d r_i^2\right)(m_2(\breve{\mathcal{S}}'_1) - 1) \leq \left(\sum_{i=2}^d r_i^2\right)(m_2(\breve{\mathcal{S}}') - 1), \tag{33}$$

and consequently

$$|\breve{\mathcal{S}}'| \leq \left(\sum_{i=2}^d r_i^2\right) m_2(\breve{\mathcal{S}}') - \sum_{i=2}^d r_i^2, \qquad \forall\, \breve{\mathcal{S}}' \subseteq \breve{\mathcal{S}}^{(0)}. \tag{34}$$

Moreover, by (32), we have $l > 6 \log(n_1) + 2\log\left(\frac{2\Pi_{i=2}^d r_i - 2\sum_{i=2}^d r_i^2}{\epsilon}\right) + 4$. Using Lemma 7, with $k = 2\Pi_{i=2}^d r_i - 2\sum_{i=2}^d r_i^2$, for each $\sum_{i=2}^d r_i^2 + 1 \leq \ell \leq \Pi_{i=2}^d r_i$, with probability at least $1 - \frac{\epsilon}{2\Pi_{i=2}^d r_i - 2\sum_{i=2}^d r_i^2}$, (30) holds for every subset of columns of $\mathcal{S}_\ell$. As a result, with probability at least $1 - \frac{\epsilon}{2}$, (30) holds for every subset of columns of $\mathcal{S}_\ell$ for $\sum_{i=2}^d r_i^2 + 1 \leq \ell \leq \Pi_{i=2}^d r_i$, simultaneously. According to Remark 8 and Lemma 8 below and by setting $r = 0$, as a subset of columns $\mathcal{S}_\ell$ of $\Omega_{(1)}$ satisfies (30), the corresponding subset of columns $\breve{\mathcal{S}}_\ell$ of the constraint matrix $\breve{\mathbf{\Omega}}$ satisfies (30) as well, $\sum_{i=2}^d r_i^2 + 1 \leq \ell \leq \Pi_{i=2}^d r_i$. Denote $\breve{\mathcal{S}}^{(1)} = \cup_{\ell=\sum_{i=2}^d r_i^2+1}^{\Pi_{i=2}^d r_i} \breve{\mathcal{S}}_\ell$. Similarly, we have

$$|\breve{\mathcal{S}}'| \leq \left(\Pi_{i=2}^d r_i - \sum_{i=2}^d r_i^2\right) m_2(\breve{\mathcal{S}}'), \qquad \forall\, \breve{\mathcal{S}}' \subseteq \breve{\mathcal{S}}^{(1)}. \tag{35}$$

For any subset of columns $\breve{\mathbf{\Omega}}''$ of $\breve{\mathbf{\Omega}}' = \breve{\mathcal{S}}^{(0)} \cup \breve{\mathcal{S}}^{(1)}$ we define $\breve{\mathcal{S}}^{(0)'}$ and $\breve{\mathcal{S}}^{(1)'}$ as the set of columns of $\breve{\mathbf{\Omega}}''$ that belong to $\breve{\mathcal{S}}^{(0)}$ and $\breve{\mathcal{S}}^{(1)}$, respectively. Observe that as $\breve{\mathbf{\Omega}}'$ has exactly $\sum_{i=1}^{\Pi_{i=2}^d r_i} |\mathcal{S}_i| = n_1 \Pi_{i=2}^d r_i - \sum_{i=2}^d r_i^2$ columns, condition (i) of Theorem 3 for $j = 1$ is satisfied. Then, by (34), (35) and the assumption that $\sum_{i=2}^d r_i^2 \leq \Pi_{i=2}^d r_i$, we have (recall $|\breve{\mathbf{\Omega}}''|$ denotes the number of columns of $\breve{\mathbf{\Omega}}''$ here)

$$\begin{aligned} |\breve{\mathbf{\Omega}}''| = |\breve{\mathcal{S}}^{(0)'}| + |\breve{\mathcal{S}}^{(1)'}| &\leq \left(\sum_{i=2}^d r_i^2\right) m_2(\breve{\mathcal{S}}^{(0)'}) - \sum_{i=2}^d r_i^2 + \\ \left(\Pi_{i=2}^d r_i - \sum_{i=2}^d r_i^2\right) m_2(\breve{\mathcal{S}}^{(1)'}) &\leq \left(\Pi_{i=2}^d r_i\right) m_2(\breve{\mathbf{\Omega}}'') - \\ \sum_{i=2}^d r_i^2 &\leq \left(\Pi_{i=2}^d r_i\right) m_2(\breve{\mathbf{\Omega}}'') - g_2(m_2(\breve{\mathbf{\Omega}}'')). \end{aligned} \tag{36}$$

Therefore, any subset of columns of $\breve{\mathbf{\Omega}}'$ satisfies (16) (condition (ii) of Theorem 3 for $j = 1$), with probability at least $1 - \epsilon$. Then, according to Theorem 3, with probability at least $1 - \epsilon$, there are only finitely many completions that fit in the sampled tensor $\mathcal{U}$, given the rank components $r_2, \ldots, r_d$. ■

Remark 6. In the case that only a subset of rank components are given, e.g, $r_{j+1}, \ldots, r_d$, we can treat the first j dimensions of the tensor as one single dimension, and therefore the above result is still applicable. In particular, assume that each column of $\widetilde{\mathbf{U}}_{(j)}$ (j-th unfolding) includes at least l observed entries, where (recall $N_j = n_1 \ldots n_j$)

$$l > 6 \log(N_j) + 2\max\left\{\log\left(\frac{2\sum_{i=j+1}^d r_i^2}{\epsilon}\right), \log\left(\frac{2\Pi_{i=j+1}^d r_i - 2\sum_{i=j+1}^d r_i^2}{\epsilon}\right)\right\} + 4. \tag{37}$$

Given that $\sum_{i=j+1}^d r_i^2 \leq \Pi_{i=j+1}^d r_i$, $\Pi_{i=j+1}^d n_i \geq N_j \Pi_{i=j+1}^d r_i - \sum_{i=j+1}^d r_i^2$, and Assumption B_j hold, with probability at least $1 - \epsilon$, there exist only finitely many completions, given the rank components $r_{j+1}, \ldots, r_d$. □

Remark 7. Assume that there exist only finitely many completions, given the rank components $r_{j+1}, \ldots, r_d$, where $j \geq 1$. Then, there exist only finitely many completions of rank $(r_1, \ldots, r_d)$. □

As in the proof of Theorem 5 we referred to Lemma 8, we need to show that if a subset of columns of $\Omega_{(1)}$ satisfies (23) or (30), the same property holds for a corresponding subset of columns of the constraint matrix $\breve{\mathbf{\Omega}}$.

Remark 8. Recall that for each column of $\Omega_{(1)}$ that has $k + r$ nonzero entries and r of them have been used to obtain $\mathbb{T}$, there exist k columns in the constraint matrix $\breve{\mathbf{\Omega}}$, each with exactly $r + 1$ nonzero entries.

Recall that as we showed after stating Assumption A_j, we can choose $\sum_{i=2}^d n_i r_i$ nonzero entries of $\Omega_{(1)}$ to obtain $\mathbb{T}$, i.e., to satisfy Assumption A_1 such that in each column of $\Omega_{(1)}$ either we choose one or zero nonzero entry. In that case, each column of $\breve{\mathbf{\Omega}}$ includes $0+1$ or $1+1$ nonzero entries. As a result, in the proof of Theorem 5, we only need Lemma 8 for

$r = 0$ and $r = 1$. However the general statement is needed to complete a proof in [34], as we explain in Remark 10.

Consequently, the matrix $\breve{\mathbf{\Omega}}'$ defined in Lemma 8 for $r = 0$ and $r = 1$ corresponds to the columns of the constraint matrix $\breve{\mathbf{\Omega}}$. □

Lemma 8. *Let r be a given nonnegative integer. Assume that there exists an $n_1 \times (n_1 - r)$ matrix Ω' composed of $n_1 - r$ columns of $\Omega_{(1)}$ such that each column of Ω' includes at least $r + 1$ nonzero entries and satisfies the following property*[2]*:*

- *Denote an $n_1 \times t$ matrix (for any $1 \leq t \leq n_1 - r$) composed of any t columns of Ω' by Ω''. Then*

$$m_2(\Omega'') - r \geq t, \tag{38}$$

where $m_2(\Omega'')$ is the number of nonzero rows of Ω''.

Then, there exists an $n_1 \times (n_1 - r)$ matrix $\breve{\mathbf{\Omega}}'$ such that: each column has exactly $r+1$ entries equal to one, and if $\breve{\mathbf{\Omega}}'(x, y) = 1$ then we have $\Omega'(x, y) = 1$. Moreover, $\breve{\mathbf{\Omega}}'$ satisfies the above-mentioned property.

Proof: Let Ω' denote a $n_1 \times (n_1 - r)$ matrix that consists of $n_1 - r$ columns of $\Omega_{(1)}$ and satisfies the mentioned property in the statement of the lemma. Note that the given property in the statement of the lemma is equivalent to the following statement

- The matrix obtained by choosing any subset of the columns of Ω' includes at least $r + t$ nonzero rows, where t is the number of the selected columns of Ω'.

Consider a bipartite graph $\mathcal{G}$ with the two sets of nodes $\mathcal{T}_1$ with $n_1 - r$ nodes corresponding to the columns and $\mathcal{T}_2$ with n_1 nodes corresponding to the rows. We add an edge between the i-th node of $\mathcal{T}_1$ and the i'-th node of $\mathcal{T}_2$ if and only if the (i, i')-th entry of Ω' is equal to one. For a set of nodes $\mathcal{S}$, let $N(\mathcal{S})$ denote the set of their neighbors. Note that according to the assumption, any subset of the columns of Ω' includes at least $r + t$ nonzero rows ($r + t$ edges in the defined graph), where t is the number of the selected columns of Ω_τ. Therefore, every subset of nodes of $\mathcal{T}_1$ has the following property

$$N(\mathcal{S}) \geq |\mathcal{S}| + r. \tag{39}$$

The statement of the lemma is equivalent to proving that there exists a spanning subgraph $\mathcal{G}'$ of $\mathcal{G}$ such that every node of $\mathcal{T}_1$ is of degree $r + 1$, and also for each subset of nodes of $\mathcal{T}_1$, (39) holds in $\mathcal{G}'$.

First consider the scenario that $r = 0$. Then, according to Hall's theorem [58], there is a perfect matching (since inequality (39) holds). Observe that a perfect matching in the described graph is exactly equivalent to what we are looking for when $r = 0$. For other values of $r > 0$, the proof is a direct result of Theorem 6 which is a **generalization** of the **Hall's theorem**. An example of such spanning subgraph for a graph with $n_1 = 6$ and $r = 1$ is shown in Figure 3. In Figure 3(a) a bipartite graph is given such that (39) holds for $r = 1$. Figure 3(b) gives an spanning subgraph of the given graph in Figure 3(a) such that degree of each node in $\mathcal{T}_1$ is equal to 2 and still (39) holds for $r = 1$. ■

[2]This property is exactly property (ii) in [34].

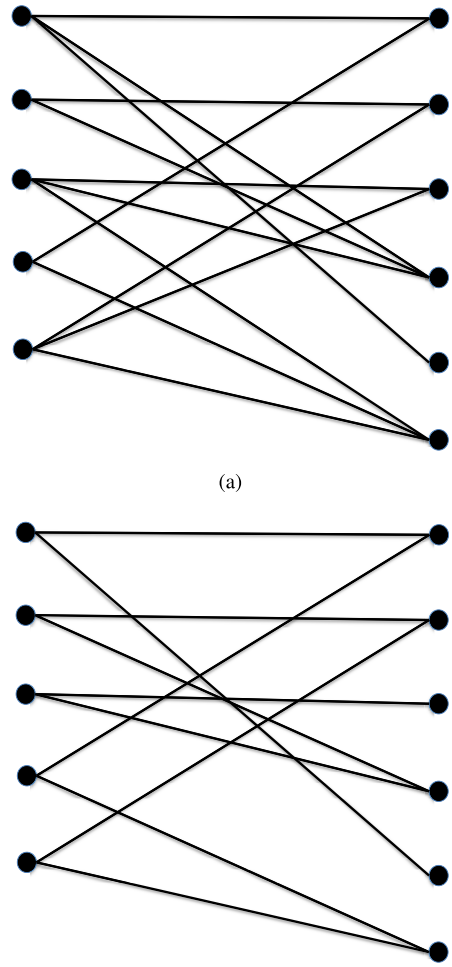

Fig. 3. An example of the corresponding bipartite graph with $r = 1$ and $n_1 = 6$. (a) The original graph. (b) The spanning subgraph.

Theorem 6. *(Generalized Hall's Theorem) Consider a bipartite graph $\mathcal{G}$ with two sets of nodes, $\mathcal{T}_1$ with x nodes and $\mathcal{T}_2$ with $x+r$ nodes, where $x, r \in \mathbb{N}$. Suppose that for each subset $\mathcal{S}$ of $\mathcal{T}_1$ the following inequality holds*

$$|N(\mathcal{S})| \geq |\mathcal{S}| + r. \tag{40}$$

Then, there exists a spanning subgraph $\mathcal{G}'$ of $\mathcal{G}$ such that every node of $\mathcal{T}_1$ is of degree $r+1$, and also for each subset of nodes of $\mathcal{T}_1$, the inequality (40) *holds in $\mathcal{G}'$ as well.*

Proof: The proof is given in Appendix B. ■

Corollary 2. *Assume that $\sum_{i=j+1}^{d} r_i^2 \leq \Pi_{i=j+1}^{d} r_i$, $\Pi_{i=j+1}^{d} n_i \geq N_j \Pi_{i=j+1}^{d} r_i - \sum_{i=j+1}^{d} r_i^2$, and Assumption B_j hold. Furthermore, assume that we observe each entry of $\mathcal{U}$ with probability p, where*

$$p > \frac{1}{N_j}\left(6\ \log\left(N_j\right) + 2\log\left(\max\left\{\frac{2\sum_{i=j+1}^{d} r_i^2}{\epsilon}, \frac{2\Pi_{i=j+1}^{d} r_i - 2\sum_{i=j+1}^{d} r_i^2}{\epsilon}\right\}\right) + 4\right) + \frac{1}{\sqrt[4]{N_j}}. \tag{41}$$

Then, with probability at least $1 - 2\epsilon$, *there are only finitely many completions for* $\mathcal{U}$, *given the rank components* $r_{j+1}, \ldots, r_d$.

Proof: The proof is straightforward given Theorem 5 and Corollary 1. ■

Remark 9. In Theorem 4 which is obtained by applying the analysis in [34], the number of needed samples per column of the i-th matricization to guarantee finite completability with high probability is on the order of $\mathcal{O}(\max\{\log(n_i), r_i\})$. Therefore, the number of needed samples in total is on the order of $\mathcal{O}(N_{-i}\max\{\log(n_i), r_i\})$. On the other hand, in Theorem 5 for the case of $d > 2$ (order of the tensor is at least 3), the number of needed samples per column for the i-th unfolding becomes $\mathcal{O}(\max\{\log(N_i), \log(r_{i+1}\ldots r_d)\})$, and the number of needed samples in total is on the order of $\mathcal{O}(\frac{N}{N_i}\max\{\log(N_i), \log(r_{i+1}\ldots r_d)\})$, if $r_{i+1}\ldots r_d < \frac{N}{N_i{}^2}$.

As an example, consider the case that $n_i = n$, $i = 1, \ldots, d$. Also, assume that $r_i = r$, $i = j+1, \ldots, d$, where $r \leq \sqrt{n}$. The matricization analysis requires $\mathcal{O}(n^{d-1}\max\{\log(n), r\})$ samples for any j. However, according to Theorem 5, if $j = \frac{d}{3}$ it requires $\mathcal{O}(n^{\frac{2d}{3}}\log(n^{\frac{d}{3}}))$ samples for finite completability. Note that under the assumption $r_{j+1}\ldots r_d < \frac{N}{N_j{}^2}$, i.e., $r^{d-j} < n^{d-2j}$, $j = \frac{d}{3}$ results in the best possible bound as $r = \sqrt{n}$ since increasing j violates the assumption. □

Remark 10. We note that the proof in [34, Theorem 3] is incomplete and Lemma 8 is needed to complete that proof.

First, observe that sampling pattern matrix Ω represents the observed entries and it has at least l nonzero elements per column. On the other hand, $\breve{\mathbf{\Omega}}$ represents the constraint matrix defined in [34] and it has exactly $(r+1)$ nonzero elements per column, where r is the given rank.

Secondly, observe that in [34, Lemma 1] each $\mathbf{\Omega}_\tau$ is a subset of the constraint matrix $\breve{\mathbf{\Omega}}$ (not a subset of Ω).

Finally, in the proof and statement of [34, Lemma 9] the authors consider a subset of Ω. However, in the proof of [34, Theorem 3] they refer to [34, Lemma 1], which considers the columns as the subsets of $\breve{\mathbf{\Omega}}$. Now, considering $\breve{\mathbf{\Omega}}$ instead of Ω in [34, Lemma 9], the equation below the equation (11) in [34] does not hold. This is because the assumption right below equation (11) in [34] that $\breve{\mathbf{\Omega}}$ has at least l nonzero elements per column is incorrect (this is a property of Ω but not necessarily $\breve{\mathbf{\Omega}}$).

As a result, a part of the proof in [34, Theorem 3] is missing which is Lemma 8 in this paper. □

V. Deterministic and Probabilistic Guarantees for Unique Completability

In this section, we first provide an example with exactly two completions to emphasize that finite completability does not necessarily result in unique completability. Then, we provide the conditions on the sampling pattern to guarantee unique completability.

Example 5. *Assume that the sampled matrix* $\mathbf{U} \in \mathbb{R}^{5\times 4}$ *is given as the incomplete matrix on the left below:*

$$\begin{array}{|c|c|c|c|}\hline 1 & & 1 & -\frac{1}{2} \\ \hline -4 & 2 & -1 & \\ \hline 0 & 1 & & 2 \\ \hline 1 & & 4 & \\ \hline & 4 & -2 & \frac{3}{2} \\ \hline \end{array} \Rightarrow \begin{array}{|c|c|c|c|}\hline 1 & -\frac{21}{32} & 1 & -\frac{1}{2} \\ \hline -4 & 2 & -1 & \frac{3}{4} \\ \hline 0 & 1 & -\frac{24}{5} & 2 \\ \hline 1 & -\frac{41}{32} & 4 & -\frac{7}{4} \\ \hline -8 & 4 & -2 & \frac{3}{2} \\ \hline \end{array} \textit{ and } \begin{array}{|c|c|c|c|}\hline 1 & -2 & 1 & -\frac{1}{2} \\ \hline -4 & 2 & -1 & -10 \\ \hline 0 & 1 & -\frac{1}{2} & 2 \\ \hline 1 & -8 & 4 & -\frac{25}{2} \\ \hline -\frac{39}{21} & 4 & -2 & \frac{3}{2} \\ \hline \end{array}$$

Moreover, assume that $rank(\mathbf{U}) = 2$. *In Appendix C, it is shown that there exist exactly two completions of* $\mathbf{U}$ *as given by the two complete matrices on the right.*

The following assumptions is a stronger version of Assumption A_j to ensure that there exists only one tuple $\mathbb{T}$ given the core tensor.

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.

Remark 11. In Lemma 1 we showed that Assumption A_j results that tuple $\mathbb{T}$ can be determined finitely. Note that Assumption A_j^+ implies Assumption A_j and therefore using only $\sum_{i=j+1}^{d}(n_i r_i)$ of the observed entries tuple $\mathbb{T}$ can be determined finitely. Moreover, observe that the rest of the sampled entries result in a set of polynomials that involve any of variables at least once and according to Note 5, tuple $\mathbb{T}$ can be determined uniquely, with probability one. □

The following lemma is a re-statement of Lemma 25 in [56] (also an adaptation of Lemma 7 in [34]) and gives the conditions under which a subset of entries of the core tensor $\mathcal{C}$ can be determined uniquely.

Lemma 9. *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: According to Lemma 3, the number of involved variables in polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}}') = \{p_1, p_2, \ldots, p_t\}$ is $t-1$ and are denoted by $\{x_1, \ldots, x_{t-1}\}$. Moreover, as mentioned in the proof of Lemma 3, $\mathcal{P}(\breve{\mathbf{\Omega}}')\backslash p_i$ is a set of algebraically independent polynomials and the number of involved variables is $t-1$, $i = 1, \ldots, t$. Consider an arbitrary variable x_1 that is involved in polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}}')$ and without loss of generality, assume that x_1 is involved in p_1.

On the other hand, as mentioned all variables $\{x_1, \ldots, x_{t-1}\}$ are involved in algebraically independent polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}}')\backslash p_1$. Hence, according to Note 5, tuple $(x_1, \ldots, x_{t-1})$

can be determined finitely. Given that one of these finitely many tuples $(x_1, \ldots, x_{t-1})$ satisfy polynomial equation p_1, according to Note 5, with probability one there does not exist another tuple $(x_1, \ldots, x_{t-1})$ among them that satisfies polynomial equation p_1. This is because with probability zero a tuple satisfy a polynomial equation in which the coefficients are chosen generically, and also the fact that the number of such tuples is finite. ■

The next theorem provides two conditions on the sampling pattern to ensure unique completability. The first one is as the same as the condition proposed in Theorem 3 for finite completability. The second condition takes advantage of finite completability and leads to the unique core tensor.

Theorem 7. *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 $t \in \{1, \ldots, n\}$ and any subtensor $\breve{\mathbf{\Omega}}'' \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times t}$ of the tensor $\breve{\mathbf{\Omega}}'$, the following inequality holds

$$\left(\Pi_{i=j+1}^{d} r_i\right) m_{j+1}(\breve{\mathbf{\Omega}}'') - g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}'')) \geq t. \quad (42)$$

(ii) For any $t' \in \{1, \ldots, n_0\}$ 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

$$\left(\Pi_{i=j+1}^{d} r_i\right) m_{j+1}(\breve{\mathbf{\Omega}}_0'') - g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}_0'')) \geq \left(\Pi_{i=j+1}^{d} r_i\right) t' - \left(\sum_{i=j+1}^{d} r_i^2\right)\left(t' - n_0 + 1\right)^+. \quad (43)$$

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

Proof: As mentioned Assumption A_j^+ results that $\mathbb{T}$ can be determined uniquely. In order to complete the proof it suffices to show the core tensor $\mathcal{C}$ can be determined uniquely as well. As defined earlier, $\mathcal{P}(\breve{\mathbf{\Omega}}')$ and $\mathcal{P}(\breve{\mathbf{\Omega}}_0')$ denote the set of polynomials obtained from the sampled entries corresponding to $\breve{\mathbf{\Omega}}'$ and $\breve{\mathbf{\Omega}}_0'$, respectively. According to Theorem 3, $\mathcal{P}(\breve{\mathbf{\Omega}}')$ results in finite completability of the sampled tensor $\mathcal{U}$, and also n algebraically independent polynomials. Hence, adding any of the polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}}_0')$ to $\mathcal{P}(\breve{\mathbf{\Omega}}')$ results in an algebraically dependent set of polynomials. Using n algebraically independent polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}}')$ and polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}}_0')$, we show all entries of core tensor can be determined uniquely.

Let p_0 denote an arbitrary polynomial in $\mathcal{P}(\breve{\mathbf{\Omega}}_0')$. Consider $\mathcal{P}'(p_0) \subset \mathcal{P}(\breve{\mathbf{\Omega}}')$ such that $\mathcal{P}'(p_0) \cup p_0$ is a minimally algebraically dependent set of polynomials. Let tuple $(x_1, \ldots, x_j)$ denote the first j coordinates of the corresponding observed entry to polynomial p_0. Then, according to Note 2, all $\Pi_{i=j+1}^{d} r_i$ entries of core tensor $\mathcal{C}$ with the first j coordinates as $(x_1, \ldots, x_j)$ are involved in polynomial p_0, and therefore they can be determined uniquely, according to Lemma 9. However, a subset of these $\Pi_{i=j+1}^{d} r_i$ entries of core tensor may be the elements of the proper structure which are known entries.

Similarly, repeating this procedure for any subtensor $\breve{\mathbf{\Omega}}_0'' \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times t'}$ of the tensor $\breve{\mathbf{\Omega}}_0'$ results in $\left(\Pi_{i=j+1}^{d} r_i\right) t'$ polynomials in terms of the $\left(\Pi_{i=j+1}^{d} r_i\right) m_{j+1}(\breve{\mathbf{\Omega}}_0'') - g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}_0''))$ unknown entries of the core tensor. Observe that according to Note 5, these polynomials are algebraically independent if

$$\left(\Pi_{i=j+1}^{d} r_i\right) m_{j+1}(\breve{\mathbf{\Omega}}_0'') - g_{j+1}(m_{j+1}(\breve{\mathbf{\Omega}}_0'')) \geq \left(\Pi_{i=j+1}^{d} r_i\right) t', \quad (44)$$

for any subtensor $\breve{\mathbf{\Omega}}_0'' \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times t'}$ of the tensor $\breve{\mathbf{\Omega}}_0'$. In order to include $\dim(\mathcal{C}) = \left(\Pi_{i=1}^{j} n_i\right)\left(\Pi_{i=j+1}^{d} r_i\right) - \left(\sum_{i=j+1}^{d} r_i^2\right)$ algebraically independent polynomials, $\breve{\mathbf{\Omega}}_0'$ should include at least n_0 columns otherwise for any subtensor $\breve{\mathbf{\Omega}}_0'' \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times t'}$ of the tensor $\breve{\mathbf{\Omega}}_0'$ we have $\dim(\mathcal{C}) > \left(\Pi_{i=j+1}^{d} r_i\right) t'$ since $t' < n_0$. Therefore, inequality (43) results that all n variables (unknown entries) of the core tensor can be determined uniquely. ■

We are also interested to obtain a lower bound on the sampling probability, which ensures the unique completability. The following lemma leads to Lemma 11 that characterizes the number of required sampled entries to ensure the unique completability with high probability.

Lemma 10. *Consider an arbitrary set $\Omega_{(1)}'$ of n_1 columns of $\Omega_{(1)}$ (first matricization of Ω). Assume that each column of $\Omega_{(1)}$ includes at least l nonzero entries, where $l > 6 \log(n_1) + 2 \log\left(\frac{n_1 k}{\epsilon}\right) + 4$. Then, with probability at least $1 - \frac{\epsilon}{k}$ every* **proper** *subset $\Omega_{(1)}''$ of columns of $\Omega_{(1)}'$ satisfies*

$$m_2(\Omega_{(1)}'') - 1 \geq t, \quad (45)$$

where t is the number of columns of $\Omega_{(1)}''$ and $m_2(\Omega_{(1)}'')$ is the number of nonzero rows of $\Omega_{(1)}''$.

Proof: According to Lemma 6, every subset $\Omega_{(1)}''$ of columns of $\Omega_{(1)}'$ with $n_1 - 1$ columns satisfies property (45) with probability at least $1 - \frac{\epsilon}{n_1 k}$. There are exactly n_1 subsets with $n_1 - 1$ columns, and therefore property (45) holds for all of them with probability at least $1 - \frac{\epsilon}{k}$. ■

Lemma 11. *Assume that $\sum_{i=2}^{d} (n_i r_i) < n_2 \ldots n_d$, $\sum_{i=2}^{d} r_i^2 \leq \Pi_{i=2}^{d} r_i$, $\Pi_{i=2}^{d} n_i \geq n_1(\Pi_{i=2}^{d} r_i + 1) - \sum_{i=2}^{d} r_i^2$, and Assumption B_1 hold. Furthermore, assume that each column of $\mathbf{U}_{(1)}$ includes at least l observed entries, where*

$$l > 6 \log(n_1) + 2 \max\left\{\log\left(\frac{\sum_{i=2}^{d} r_i^2}{\epsilon}\right), \log\left(\frac{\Pi_{i=2}^{d} r_i - \sum_{i=2}^{d} r_i^2}{\epsilon}\right), \log\left(\frac{n_1}{\epsilon}\right)\right\} + 8. \quad (46)$$

Then, with probability at least $1 - \epsilon$, there exists only one completion of rank $(r_1, \cdots, r_d)$.

Proof: Observe that since $\Pi_{i=2}^d n_i \geq n_1(\Pi_{i=2}^d r_i + 1) - \sum_{i=2}^d r_i^2$ there exist $n_1\left(\Pi_{i=2}^d r_i\right) - \left(\sum_{i=2}^d r_i^2\right)$ and n_1 disjoint columns in $\Omega_{(1)}$ denoted by $\breve{\boldsymbol{\Omega}}'$ and $\breve{\boldsymbol{\Omega}}_0'$, respectively. In order to show unique completability, it suffices to show $\breve{\boldsymbol{\Omega}}'$ and $\breve{\boldsymbol{\Omega}}_0'$ satisfy properties (i) and (ii) in the statement of Theorem 7, respectively, with probability at least $1 - \frac{\epsilon}{2}$.

Note that having (46), it is easy to see that (32) holds with ϵ replaced by $\frac{\epsilon}{2}$. Therefore, according to Theorem 5, $\breve{\boldsymbol{\Omega}}'$ satisfies property (i) in the statement of Theorem 7 with probability at least $1 - \frac{\epsilon}{2}$.

On the other hand, $\sum_{i=2}^d r_i^2 \leq \Pi_{i=2}^d r_i$ and according to Lemma 10 (with $k = 2$), with probability at least $1 - \frac{\epsilon}{2}$, any subset of columns of $\breve{\boldsymbol{\Omega}}_0'$ satisfies (45). For any **proper** subset of column $\breve{\boldsymbol{\Omega}}_0''$ of $\breve{\boldsymbol{\Omega}}_0'$, we have $m_2(\breve{\boldsymbol{\Omega}}_0'') - 1 \geq t'$ or equivalently (t' is the number of columns of $\breve{\boldsymbol{\Omega}}_0''$)

$$\left(\Pi_{i=j+1}^d r_i\right) m_{j+1}(\breve{\boldsymbol{\Omega}}_0'') - \Pi_{i=2}^d r_i \geq \left(\Pi_{i=j+1}^d r_i\right) t', \quad (47)$$

which results in (43) as $g_{j+1}(m_{j+1}(\breve{\boldsymbol{\Omega}}_0'')) \leq \sum_{i=2}^d r_i^2 \leq \Pi_{i=2}^d r_i$ and $\left(t' - n_0 + 1\right)^+ = 0$. Hence, in order to show property (ii) in the statement of Theorem 7 holds with probability at least $1 - \frac{\epsilon}{2}$ (which completes the proof) we only need to show that (43) holds when $\breve{\boldsymbol{\Omega}}_0'' = \breve{\boldsymbol{\Omega}}_0'$.

Note that $m_{j+1}(\breve{\boldsymbol{\Omega}}_0'') \leq m_{j+1}(\breve{\boldsymbol{\Omega}}_0')$. Therefore, as $m_2(\breve{\boldsymbol{\Omega}}_0'') - 1 \geq t'$ holds for any proper subset of column $\breve{\boldsymbol{\Omega}}_0''$ of $\breve{\boldsymbol{\Omega}}_0'$, we have $m_{j+1}(\breve{\boldsymbol{\Omega}}_0') \geq n_1$ (by setting $t' = n_1 - 1$). Then, (43) holds as $g_{j+1}(m_{j+1}(\breve{\boldsymbol{\Omega}}_0')) \leq \sum_{i=2}^d r_i^2$ and $\left(t' - n_0 + 1\right)^+ = 1$. ■

Finally, in the following, we use Corollary 1 and the above lemma to propose a lower bound on the sampling probability to guarantee the unique completability with high probability.

Corollary 3. *Assume that* $\sum_{i=j+1}^d r_i^2 \leq \Pi_{i=j+1}^d r_i$, $\Pi_{i=j+1}^d n_i \geq N_j(\Pi_{i=j+1}^d r_i + 1) - \sum_{i=j+1}^d r_i^2$, *and Assumption* B_j *hold. Furthermore, assume that we observe each entry of the tensor* $\mathcal{U}$ *with probability* p, *where*

$$p > \frac{1}{N_j}\left(6\ \log\left(N_j\right) + 2\log\left(\max\left\{\frac{\sum_{i=j+1}^d r_i^2}{\epsilon}, \frac{\Pi_{i=j+1}^d r_i - \sum_{i=j+1}^d r_i^2}{\epsilon}, \frac{N_j}{\epsilon}\right\}\right) + 8\right) + \frac{1}{\sqrt[4]{N_j}}. \quad (48)$$

Then, with probability at least $1 - 2\epsilon$, *there exists only one completion for* $\mathcal{U}$, *given the rank components* $r_{j+1}, \ldots, r_d$.

Proof: The proof is straightforward given Lemma 11 and Corollary 1. ■

VI. Numerical Comparisons

In order to show the advantage of our proposed method over matrix analysis, we compare the lower bound on the sampling probability obtained by matricization analysis with the bound obtained by tensor analysis.

In the first example, numerical comparisons are performed on a 4th-order tensor $\mathcal{U} \in \mathbb{R}^{900\times900\times900\times900}$. Figure 4 plots the bounds given in (21) (Grassmannian manifold) and (41) (Tucker manifold) for finite completability, where $r_1 = \cdots = r_4 = r$, and $\epsilon = 0.0001$. As the second example, we consider a 6th-order tensor $\mathcal{U} \in \mathbb{R}^{900\times900\times900\times900\times900\times900}$. Figure 5 plots the bounds given in (21) (Grassmannian manifold) and (41) (Tucker manifold) for finite completability, where $r_3 = r_4 = r_5 = r_4 = r$, and $\epsilon = 0.0001$. In this example, a significant reduction in the sampling probability is seen for the tensor model.

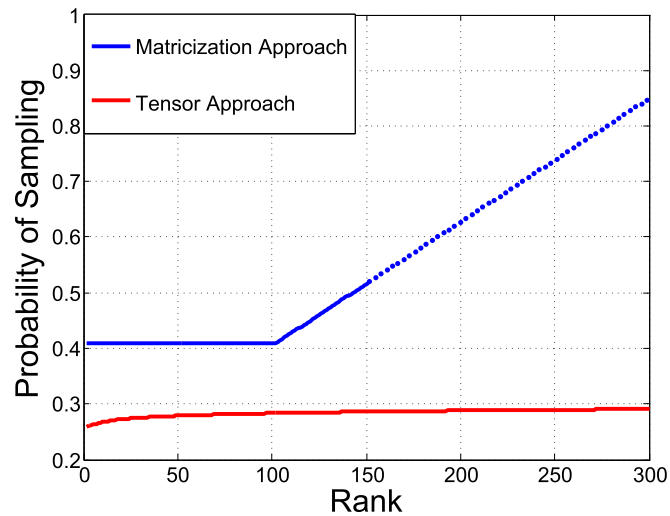

Fig. 4. Lower bounds on sampling probability for a 4th-order tensor.

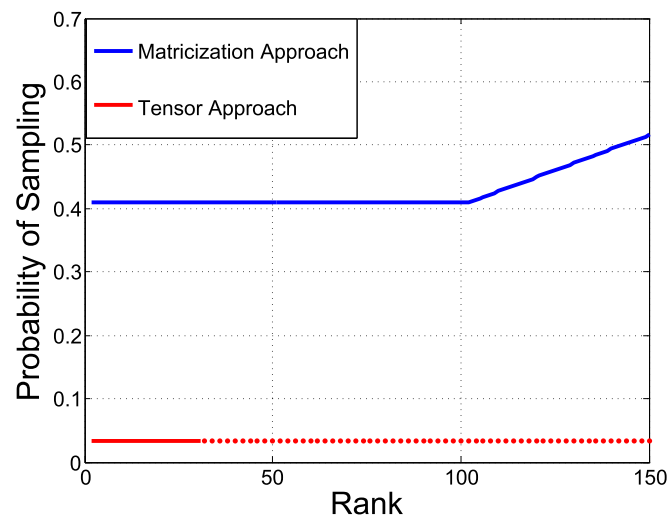

Fig. 5. Lower bounds on sampling probability for a 6th-order tensor.

We have the following observations:

- In Figure 4, the bound obtained through the analysis on Tucker manifold is less than the bound obtained through the low-order analysis on Grassmannian manifold. This improvement significantly increases as the value of the rank increases.
- In Figure 4, the restriction $r \leq \frac{n}{6}$ in the analysis on Grassmannian manifold makes the bound valid only for $r \leq 150$ (and that is why the corresponding curve is dotted for $r > 150$ in Figure 4). On the other hand, the restriction $r(d-1) \leq n$ in our proposed approach ensures the validity of the bound for $r \leq 300$. Similarly, in Figure 5,

the restriction $\Pi_{i=j+1}^{d} n_i \geq N_j \Pi_{i=j+1}^{d} r_i - \sum_{i=j+1}^{d} r_i^2$ in the analysis on Tucker manifold makes the bound valid only for $r \leq 30$.

- Since the bounds for finite completability (48) and unique completability (41) result in almost the same curves in the above examples, we only plot bounds for finite completability in the figures. In general, the main difference between (41) and (48) is an additional term $\log\left(\frac{N_j}{\epsilon}\right)$ in the second term inside of maximum operator.

VII. Conclusions

This paper aims to find fundamental conditions on the sampling pattern for finite completability of a low-rank partially sampled tensor. 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 using Bernstein's theorem and analysis on Tucker manifold, a relationship between a geometric pattern on the sampled entries and the number of algebraically independent polynomials among all of the defined polynomials is characterized. Moreover, an extension to Hall's theorem in graph theory is provided which is key to obtaining probabilistic conditions for finite and unique completabilities. Using these developed tools, we have addressed three problems in this paper: (i) Characterizing the necessary and sufficient conditions on the sampling pattern to have finitely many tensor completions for the given rank, (ii) Characterizing conditions on the sampling pattern to ensure that there is exactly one completion for the given rank, (iii) Lower bounding the sampling probability such that the conditions in Problems (i) and (ii) are satisfied with high probability.

Finally, through numerical examples it is seen that our proposed analysis on Tucker manifold for finite tensor completability leads to a much lower sampling rate than the matricization approach that is based on analysis on Grassmannian manifold.

Appendix A
Proof of Lemma 4

As mentioned before, finiteness of the number of completions given $(r_{j+1}, \ldots, r_d)$ results finiteness of the number of completions given $(r_j, \ldots, r_d)$. We are interested to show this statement holds in Theorem 3, i.e, satisfying the properties (i) and (ii) for j in the statement of Theorem 3 results in satisfying the properties (i) and (ii) for $j-1$. Before we present this result, the following notation is introduced for the simplicity.

Definition 6. *Let set $\mathcal{S}$ be a set of d-tuples such that each element of $\mathcal{S}$ consists of d natural numbers $(x_1, \ldots, x_d)$. Define $f_j(\mathcal{S})$ as the number of different tuples of the first j components of the elements of $\mathcal{S}$.*

For example, let $\mathcal{S} = \{(2,2,1), (2,3,3), (2,2,2), (1,2,1), (2,3,1)\}$. Then, according to the definition $f_2(\mathcal{S}) = |\{(2,2), (2,3), (1,2)\}| = 3$.

Now, we are ready to provide the proof of lemma. It is easy to see that Lemma 4 is exactly equivalent to the following statement:

Assume that we sample a tensor $\mathcal{U} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$. Suppose that there exist $n = \left(\Pi_{i=1}^{j} n_i\right)\left(\Pi_{i=j+1}^{d} r_i\right) - \left(\sum_{i=j+1}^{d} r_i^2\right)$ of the sampled entries such that any subset of them (called $\mathcal{S}$) with t observed entires (for any $1 \leq t \leq n$) satisfies the following inequality

$$\left(\Pi_{i=j+1}^{d} r_i\right) f_j(\mathcal{S}) - g_{j+1}(f_j(\mathcal{S})) \geq t. \tag{49}$$

Then, there exist $n' = \left(\Pi_{i=1}^{j-1} n_i\right)\left(\Pi_{i=j}^{d} r_i\right) - \left(\sum_{i=j}^{d} r_i^2\right)$of the sampled entries such that any subset of them (called $\mathcal{S}'$) with t observed entires (for any $1 \leq t \leq n'$) satisfies the following inequality

$$\left(\Pi_{i=j}^{d} r_i\right) f_{j-1}(\mathcal{S}') - g_j(f_{j-1}(\mathcal{S}')) \geq t. \tag{50}$$

Therefore, it suffices to show the above statement. For $r_j = 1$, the statement can be easily verified, and therefore we consider the case of $r_j > 1$ in the rest of proof. Also recall the assumption $n_j > \sum_{i=1}^{d} r_i$. We partition all of the sampled entries into n_j groups such that group i includes all of the sampled entries that j-th component of their coordinate is equal to i. Let $\mathcal{S}_i$ denote the i-th group. Observe that every subset of the sampled entries of each group satisfies inequality (49). Moreover, according to pigeonhole principle we know that between the defined n_j groups there exist r_j groups (without loss of generality assume $\mathcal{S}_1, \ldots, \mathcal{S}_{r_j}$) such that $\sum_{i=1}^{r_j} |\mathcal{S}_i| \geq \frac{r_j}{n_j} n$.

Since $\frac{r_j}{n_j} n = \left(\Pi_{i=1}^{j-1} n_i\right)\left(\Pi_{i=j}^{d} r_i\right) - \frac{r_j}{n_j}\left(\sum_{i=j+1}^{d} r_i^2\right) \geq n'$, we only need to show that any subset of the elements of the groups $\mathcal{S}_1, \ldots, \mathcal{S}_{r_j}$ satisfies inequality (50). Consider an arbitrary subset of the elements of groups $\mathcal{S}_1, \ldots, \mathcal{S}_{r_j}$ and denote it by $\mathcal{S}'$. Assume that $\mathcal{S}'_i$ denotes the elements of $\mathcal{S}'$ that belong in $\mathcal{S}_i$ for $1 \leq i \leq r_j$. Recall that every subset of the sampled entries of each group satisfies inequality (49). Moreover, due to the fact that in each group the j-th component of the coordinates are the same, for each subset of the entries of a group $\mathcal{S}_i$ we have $f_j(\mathcal{S}'_i) = f_{j-1}(\mathcal{S}'_i)$. As a result, we have

$$\left(\Pi_{i=j+1}^{d} r_i\right) f_{j-1}(\mathcal{S}'_i) - g_{j+1}(f_{j-1}(\mathcal{S}'_i)) \geq |\mathcal{S}'_i|. \tag{51}$$

Without loss of generality, assume $f_j(\mathcal{S}'_1) \geq f_j(\mathcal{S}'_2) \geq \ldots \geq f_j(\mathcal{S}'_{r_j})$. Also, observe that for each $1 \leq i \leq r_j$ we have $f_j(\mathcal{S}'_i) \leq f_j(\mathcal{S}')$. Therefore, by adding up the above inequalities for $1 \leq i \leq r_j$ we have

$$\left(\Pi_{i=j}^{d} r_i\right) f_{j-1}(\mathcal{S}'_1) - g_{j+1}(f_{j-1}(\mathcal{S}'_1)) r_j \geq |\mathcal{S}'|. \tag{52}$$

For the case that $1 < r_j < \sum_{i=j+1}^{d} r_i^2$ holds, we can see that $r_j\left(\sum_{i=j+1}^{d} r_i^2\right) \geq \sum_{i=j}^{d} r_i^2$ also holds, and therefore using (52) we can obtain

$$\left(\Pi_{i=j}^{d} r_i\right) f_{j-1}(\mathcal{S}') - g_j(f_{j-1}(\mathcal{S}')) \geq |\mathcal{S}'|, \tag{53}$$

which completes the proof. If $r_j \geq \sum_{i=j+1}^{d} r_i^2$ similarly using the above inequality and by ignoring r_j^2 of entries proof can be completed.

APPENDIX B
PROOF OF THEOREM 6

The proof of this theorem is based on strong induction on x. For $x = 1$ the theorem is easy to verify. Assume that the statement of the theorem holds for $1 \leq x \leq x_0$. Then, it suffices to show that the statement also holds for the case that $x = x_0 + 1$. There are two different scenarios that we need to show the statement holds for separately.

Scenario 1. *There exists a proper and nonempty subset of nodes $\mathcal{S}_1$ of $\mathcal{T}_1$ such that $|N(\mathcal{S}_1)| = |\mathcal{S}_1| + r$:*

Consider the induced subgraph of $\mathcal{G}$ with the set of nodes $\mathcal{S}_1 \cup N(\mathcal{S}_1)$ and denote it by $\mathcal{G}_1$. Induction hypothesis results in a spanning subgraph $\mathcal{G}_1'$ of $\mathcal{G}_1$ that satisfies the desired properties in the statement of the theorem for the corresponding subgraph $\mathcal{G}_1$.

Now, consider the induced subgraph of $\mathcal{G}$ with the set of nodes $\{\mathcal{T}_1 \cup \mathcal{T}_2\} \setminus \{\mathcal{S}_1\}$ and denote it by $\mathcal{G}_2$. Since $|N(\mathcal{S}_1)| = |\mathcal{S}_1| + r$, we conclude that for each subset of nodes of $\{\mathcal{T}_1\}\setminus\{\mathcal{S}_1\}$ we have $|N(\mathcal{S}) \cap \{\mathcal{T}_2 \setminus N(\mathcal{S}_1)\}| \geq |\mathcal{S}|$. The reason is that if we choose a set of nodes including the members of $\mathcal{S} \subset \{\mathcal{T}_1\}\setminus\{\mathcal{S}_1\}$ plus all nodes in $\mathcal{S}_1$ and use the assumption in the statement of the theorem, it results that $|N(\mathcal{S}) \cap \{\mathcal{T}_2 \setminus N(\mathcal{S}_1)\}| \geq |\mathcal{S}|$. Moreover, induction hypothesis results that there exists a spanning subgraph that satisfies the desired properties in the statement of the theorem for the corresponding subgraph $\mathcal{G}_2$. Now, Lemma 12 results that there exists a spanning subgraph $\mathcal{G}_2'$ of the graph $\mathcal{G}_2$ so that every node of $\mathcal{T}_1 \setminus \mathcal{S}_1$ is of degree $r+1$, and also for each subset of nodes of $\mathcal{T}_1\setminus\mathcal{S}_1$, the inequality (40) holds and in addition, $\mathcal{G}_2'$ includes a perfect matching between the nodes in $\mathcal{T}_1 \setminus \mathcal{S}_1$ and $\mathcal{T}_2 \setminus N(\mathcal{S}_1)$.

Now, consider a spanning subgraph $\mathcal{G}'$ of the graph $\mathcal{G}$ that only includes all edges of $\mathcal{G}_1'$ and $\mathcal{G}_2'$. We can easily observe that every node of $\mathcal{T}_1$ is of degree $r + 1$. Moreover, for each subset of nodes of $\mathcal{S}_1$ the inequality (40) holds, and also for each subset of nodes of $\mathcal{T}_1 \setminus \mathcal{S}_1$ the inequality (40) holds. Now, consider a subset of nodes $\mathcal{S}$ of $\mathcal{T}_1$. It suffices to show the inequality (40) holds for $\mathcal{S}$.

Define $\mathcal{S}' = \mathcal{S} \cap \mathcal{S}_1$ and $\mathcal{S}'' = \mathcal{S} \cap \{\mathcal{T}_1 \setminus \mathcal{S}_1\}$. Since for each subset of nodes of $\mathcal{S}_1$ the inequality (40) holds, we have $|N(\mathcal{S}) \cap N(\mathcal{S}_1)| \geq |\mathcal{S}'| + r$. On the other hand, $\mathcal{G}'$ includes a perfect matching between the nodes in $\mathcal{T}_1\setminus\mathcal{S}_1$ and $\mathcal{T}_2 \setminus N(\mathcal{S}_1)$, and therefore $|N(\mathcal{S}) \cap \{\mathcal{T}_2 \setminus N(\mathcal{S}_1)\}| \geq |\mathcal{S}''|$. As a result, $|N(\mathcal{S})| \geq |\mathcal{S}'| + r + |\mathcal{S}''| = |\mathcal{S}| + r$ and the proof is complete for this scenario.

Scenario 2: *For any proper and nonempty subset of nodes of $\mathcal{T}_1$ we have $|N(\mathcal{S})| \geq |\mathcal{S}| + r + 1$:*

Consider an arbitrary node v_1 and observe that $|N(v_1)| \geq r + 1$. Define $\mathcal{S}_1 = \mathcal{T}_1 \setminus \{v_1\}$ and let $\mathcal{G}_1$ denote the induced subgraph of $\mathcal{G}$ with set of nodes $\mathcal{S}_1 \cup \mathcal{T}_2$ (which is all nodes of graph $\mathcal{G}$ except for the node v_1). Induction hypothesis results that there exists a spanning subgraph $\mathcal{G}_1'$ of the graph $\mathcal{G}_1$ such that every node of $\mathcal{S}_1$ is of degree $r + 1$, and also for each subset of nodes of $\mathcal{S}_1$, the inequality (40) holds. Now, consider a spanning subgraph $\mathcal{G}'$ of the graph $\mathcal{G}$ that includes only all edges of $\mathcal{G}_1'$, and also $r + 1$ random edges among all edges that are connected to v_1. It is clear that $\mathcal{G}'$ satisfies all of the desired properties mentioned in the statement of the theorem.

In order to complete the proof of Theorem 6, we need the following lemma as it was mentioned before.

Lemma 12. *Consider a bipartite graph $\mathcal{G}$ with the two sets of nodes, $\mathcal{T}_1$ with x nodes and $\mathcal{T}_2$ with $x+r+y$ nodes, where $x, y, r \in \mathbb{N}$. Suppose that for each subset of the nodes of $\mathcal{T}_1$ we have*

$$|N(\mathcal{S})| \geq |\mathcal{S}| + r. \tag{54}$$

Moreover, assume that there exists a subset of nodes $\mathcal{S}_0$ of $\mathcal{T}_2$ such that $|\mathcal{S}_0| = x$, and also for every subset of nodes $\mathcal{S}$ of $\mathcal{T}_1$ we have $|N(\mathcal{S}) \cap \mathcal{S}_0| \geq |\mathcal{S}|$. Assume that there exists a spanning subgraph $\mathcal{G}_0$ of $\mathcal{G}$ such that every node of $\mathcal{T}_1$ is of degree $r + 1$, and also for each subset of nodes of $\mathcal{T}_1$, the inequality (54) *holds. Then, there exists a spanning subgraph $\mathcal{G}'$ of $\mathcal{G}$ such that every node of $\mathcal{T}_1$ is of degree $r+1$, and also for each subset of nodes of $\mathcal{T}_1$, the inequality* (54) *holds and in addition, $\mathcal{G}'$ includes a perfect matching between the nodes in $\mathcal{T}_1$ and $\mathcal{S}_0$.*

Proof: We prove the lemma using strong induction on x. For $x = 1$ the lemma is easy to verify. Assume that the statement of lemma holds for $1 \leq x \leq x_0$. Then, we only need to show that lemma holds for the case that $x = x_0 + 1$. There are two scenarios and we prove the statement for these two scenarios separately as follows.

Scenario 1: *There exists a proper and nonempty subset of nodes $\mathcal{S}_1$ of $\mathcal{T}_1$ such that $|N(\mathcal{S}_1) \cap \mathcal{S}_0| = |\mathcal{S}_1|$:*

For simplicity in notation, define $\mathcal{S}_1' \triangleq \mathcal{T}_1 \setminus \mathcal{S}_1$. In this scenario, we use the induction hypothesis for $\mathcal{S}_1$ and $\mathcal{S}_1'$ separately (since $|\mathcal{S}_1| \leq x_0$ and $|\mathcal{S}_1'| \leq x_0$). Define the sets $\mathcal{S}_0' = N(\mathcal{S}_1) \cap \mathcal{S}_0$ and $\mathcal{S}_0'' = \mathcal{S}_0 \setminus \mathcal{S}_0'$. Consider the induced subgraph $\mathcal{G}_1$ of the graph $\mathcal{G}$ where the set of vertices of $\mathcal{G}_1$ is $\mathcal{T}_2 \cup \mathcal{S}_1$. Assumption of the lemma results that there exists a spanning subgraph of $\mathcal{G}_1$ such that every node of $\mathcal{S}_1$ is of degree $r + 1$, and also for each subset of nodes of $\mathcal{S}_1$, the inequality (54) holds (by considering all the edges of the induced subgraph with vertices $\mathcal{T}_2 \cup \mathcal{S}_1$ that also exist in $\mathcal{G}_0$). Induction hypothesis results that there exists a spanning subgraph $\mathcal{G}_0'$ of $\mathcal{G}_1$ such that every node of $\mathcal{S}_1$ is of degree $r+1$, and also for each subset of nodes of $\mathcal{S}_1$, the inequality (54) holds and in addition, $\mathcal{G}_0'$ includes a perfect matching between nodes of $\mathcal{S}_1$ and nodes of $\mathcal{S}_0'$.

Consider the induced subgraph $\mathcal{G}_1'$ of the graph $\mathcal{G}$ where the set of vertices of $\mathcal{G}_2$ is $\mathcal{T}_2 \cup \mathcal{S}_1'$. Again, assumption of the lemma results that there exists a spanning subgraph of $\mathcal{G}_1'$ such that every node of $\mathcal{S}_1'$ is of degree $r + 1$, and also for each subset of nodes of $\mathcal{S}_1'$, the inequality (54) holds (by considering all the edges of the induced subgraph with vertices $\mathcal{T}_2 \cup \mathcal{S}_1'$ that also exist in $\mathcal{G}_0$). Also, observe that since $|N(\mathcal{S}_1) \cap \mathcal{S}_0| = |\mathcal{S}_1|$, for every subset of nodes $\mathcal{S}$ of $\mathcal{S}_1'$ we have $|N(\mathcal{S}) \cap \mathcal{S}_0''| \geq |\mathcal{S}|$. As a result, induction hypothesis results that there exists a spanning subgraph $\mathcal{G}_0''$ of $\mathcal{G}_1'$ such that every node of $\mathcal{S}_1'$ is of degree $r + 1$, and also for each subset of nodes of $\mathcal{S}_1'$, the inequality (54) holds. In addition, $\mathcal{G}_0''$ includes a perfect matching between nodes of $\mathcal{S}_1'$ and nodes of $\mathcal{S}_0''$.

Now, consider a spanning subgraph $\mathcal{G}'$ of the graph $\mathcal{G}$ that only includes all edges of $\mathcal{G}_0'$ and $\mathcal{G}_0''$. It can be verified that

$\mathcal{G}'$ satisfies all of the desired properties in the statement of the lemma and therefore the proof is complete for this case.

Scenario 2: *For any proper and nonempty subset of nodes of $\mathcal{T}_1$ we have $|N(\mathcal{S}) \cap \mathcal{S}_0| \geq |\mathcal{S}| + 1$:*

In this case, consider an arbitrary vertex of $\mathcal{T}_1$ and denote it by v_0, and also define $\mathcal{S}_1 = \mathcal{T}_1 \backslash \{v_0\}$. Hence, we have $|N(\mathcal{S}_1) \cap \mathcal{S}_0| \geq |\mathcal{S}_1| + 1$ and since for any $\mathcal{S}$ we have $|N(\mathcal{S}) \cap \mathcal{S}_0| \leq |\mathcal{S}_0|$ we conclude that $|N(\mathcal{S}_1) \cap \mathcal{S}_0| = |\mathcal{S}_1| + 1$. Also, according to the assumptions of the lemma we know $N(v_0) \geq r + 1$ and $|N(v_0) \cap \mathcal{S}_0| \geq 1$. We choose an arbitrary node among the nodes in $|N(v_0) \cap \mathcal{S}_0|$ and denote it by u_0. Construct a graph $\mathcal{G}_0''$ with nodes $\mathcal{T}_2 \cup \{v_0\}$ and $r + 1$ edges among edges that are connected to v_0 including the edge between v_0 and u_0, i.e., (v_0, u_0).

Now, consider the induced subgraph $\mathcal{G}_1$ of the graph $\mathcal{G}$ where the set of vertices of $\mathcal{G}_1$ is $\mathcal{T}_2 \cup \mathcal{S}_1$, i.e., all nodes in $\mathcal{G}$ except for the node v_0. Assumption of the lemma results that there exists a spanning subgraph of $\mathcal{G}_1$ such that every node of $\mathcal{S}_1$ is of degree $r + 1$, and also for each subset of nodes of $\mathcal{S}_1$, the inequality (54) holds (by considering all edges of the induced subgraph with vertices $\mathcal{T}_2 \cup \mathcal{S}_1$ that also exist in $\mathcal{G}_0$). Induction hypothesis results that there exists a spanning subgraph $\mathcal{G}_0'$ of $\mathcal{G}_1$ such that every node of $\mathcal{S}_1$ is of degree $r+1$, and also for each subset of nodes of $\mathcal{S}_1$, the inequality (54) holds and in addition, $\mathcal{G}_0'$ includes a perfect matching between nodes of $\mathcal{S}_1$ and nodes of $\mathcal{S}_0 \setminus \{u_0\}$.

Finally, consider a spanning subgraph $\mathcal{G}'$ of the graph $\mathcal{G}$ that only includes all edges of $\mathcal{G}_0'$ and $\mathcal{G}_0''$. The constructed graph $\mathcal{G}'$ satisfies all of the mentioned properties in the lemma. ■

Appendix C
Proof of Claim in Example 5

In the following we show that given the rank-2 sampled matrix in Example 5, there are exactly two completions. Note that the following decomposition always holds for some values of $r_1, r_2, \ldots, r_6$ and $x_1, x_2, \ldots, x_8$:

$$\begin{bmatrix} 1 & & 1 & -\frac{1}{2} \\ -4 & 2 & -1 & \\ 0 & 1 & & 2 \\ 1 & & 4 & \\ & 4 & -2 & \frac{3}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ r_1 & r_2 \\ r_3 & r_4 \\ r_5 & r_6 \end{bmatrix} \times \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ x_5 & x_6 & x_7 & x_8 \end{bmatrix}$$

Note that the 2×2 identity matrix in the above decomposition represents canonical basis defined in Definition 12. The first two rows of $\mathbf{U}$ in the above decomposition result the following:

$$1 = x_1, \tag{55a}$$
$$1 = x_3, \tag{55b}$$
$$-\frac{1}{2} = x_4, \tag{55c}$$
$$-4 = x_5, \tag{55d}$$
$$2 = x_6, \tag{55e}$$
$$-1 = x_7. \tag{55f}$$

Then, the decomposition can be simplified as

$$\begin{bmatrix} 0 & 1 & & 2 \\ 1 & & 4 & \\ & 4 & -2 & \frac{3}{2} \end{bmatrix} = \begin{bmatrix} r_1 & r_2 \\ r_3 & r_4 \\ r_5 & r_6 \end{bmatrix} \times \begin{bmatrix} 1 & x_2 & 1 & -\frac{1}{2} \\ -4 & 2 & -1 & x_8 \end{bmatrix}$$

Therefore, we have following system of equations

$$0 = r_1 - 4r_2, \tag{56a}$$
$$1 = x_2 r_1 + 2r_2, \tag{56b}$$
$$2 = -\frac{1}{2} r_1 + x_8 r_2, \tag{56c}$$
$$1 = r_3 - 4r_4, \tag{56d}$$
$$4 = r_3 - r_4, \tag{56e}$$
$$4 = x_2 r_5 + 2r_6, \tag{56f}$$
$$-2 = r_5 - r_6, \tag{56g}$$
$$\frac{3}{2} = -\frac{1}{2} r_5 + x_8 r_6. \tag{56h}$$

Observe that $r_3 = 5$ and $r_4 = 1$ can be determined uniquely from (56d) and (56e). Note that $r_1 = 4r_2$ and $r_5 = (r_6 - 2)$ can be concluded from (56a) and (56g), respectively. Then, by substituting $r_1 \leftarrow 4r_2$ in (56b) and (56c) and substituting $r_5 \leftarrow (r_6 - 2)$ in (56f) and (56h) we obtain:

$$1 = 4x_2 r_2 + 2r_2, \tag{57a}$$
$$2 = -2r_2 + x_8 r_2, \tag{57b}$$
$$4 = x_2(r_6 - 2) + 2r_6, \tag{57c}$$
$$\frac{3}{2} = -\frac{1}{2}(r_6 - 2) + x_8 r_6. \tag{57d}$$

Similarly, $r_2 = \frac{2}{x_8 - 2}$ and $r_6 = \frac{1}{2x_8 - 1}$ can be concluded from (57b) and (57d), respectively. Hence, substituting $r_2 \leftarrow \frac{2}{x_8 - 2}$, $r_6 \leftarrow \frac{1}{2x_8 - 1}$ in (57a) and (57c) results in the following system of equations

$$x_8 = 8x_2 + 6, \tag{58a}$$
$$8x_8 - 4 = x_2 - 2x_2(2x_8 - 1) + 2. \tag{58b}$$

Finally, (58a) results $x_8 = 8x_2 + 6$, and therefore substituting $x_8 \leftarrow 8x_2 + 6$ in (58b) results

$$8(8x_2 + 6) - 4 = x_2 - 2x_2(2(8x_2 + 6) - 1) + 2, \tag{59}$$

which can be simplified as

$$32x_2^2 + 85x_2 + 42 = 0. \tag{60}$$

Therefore, $x_2 \in \{-2, -\frac{21}{32}\}$. Given x_2, all other variables can be determined uniquely recursively as above, and the following completions of $\mathbf{U}$ can be obtained as the only possible rank-2 completions of $\mathbf{U}$.

$$\begin{bmatrix} 1 & -\frac{21}{32} & 1 & -\frac{1}{2} \\ -4 & 2 & -1 & \frac{3}{4} \\ 0 & 1 & -\frac{24}{5} & 2 \\ 1 & -\frac{41}{32} & 4 & -\frac{7}{4} \\ -8 & 4 & -2 & \frac{3}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -\frac{32}{5} & -\frac{8}{5} \\ 5 & 1 \\ 0 & 2 \end{bmatrix} \times \begin{bmatrix} 1 & -\frac{21}{32} & 1 & -\frac{1}{2} \\ -4 & 2 & -1 & \frac{3}{4} \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & -2 & 1 & -\frac{1}{2} \\ -4 & 2 & -1 & -10 \\ 0 & 1 & -\frac{1}{2} & 2 \\ 1 & -8 & 4 & -\frac{25}{2} \\ -\frac{39}{21} & 4 & -2 & \frac{3}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -\frac{2}{3} & -\frac{1}{6} \\ 5 & 1 \\ -\frac{43}{21} & -\frac{1}{21} \end{bmatrix} \times \begin{bmatrix} 1 & -2 & 1 & -\frac{1}{2} \\ -4 & 2 & -1 & -10 \end{bmatrix}$$

REFERENCES

[1] E. J. Candès, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM Rev.*, 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 *Proc. CVPR*, Jun. 2010, pp. 1791–1798.

[3] L. Eldén, *Matrix Methods in Data Mining and Pattern Recognition*, vol. 4. Philadelphia, PA, USA: SIAM, 2007.

[4] N. J. A. Harvey, D. R. Karger, and K. Murota, "Deterministic network coding by matrix completion," in *Proc. 16th Annu. ACM-SIAM Symp. Discrete Algorithms*, Jan. 2005, pp. 489–498.

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

[6] N. D. Sidiropoulos and A. Kyrillidis, "Multi-way compressed sensing for sparse low-rank tensors," *IEEE Signal Process. Lett.*, vol. 19, no. 11, pp. 757–760, Nov. 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, an. 2011, Art. no. 025010.

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

[9] X.-Y. Liu, S. Aeron, V. Aggarwal, and X. Wang, "Low-tubal-rank tensor completion using alternating minimization," *Proc. SPIE*, vol. 9848, May 2016, Art. no. 984809. [Online]. Available:https://www.spiedigitallibrary.org/conference-proceedings-of-spie/9848/1/Low-tubal-ranktensor-completion-using-alternating-minimization/10.1117/12.2224039.full?SSO=1

[10] 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.

[11] G. Ely *et al.*, "5D and 4D pre-stack seismic data completion using tensor nuclear norm (TNN)," in *Proc. SEG Annu. Meeting*, Sep. 2013, pp. 3639–3644.

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

[13] 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 *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2016, pp. 2529–2533.

[14] X.-Y. Liu, S. Aeron, V. Aggarwal, X. Wang, and M.-Y. Wu, "Adaptive sampling of RF fingerprints for fine-grained indoor localization," *IEEE Trans. Mobile Comput.*, vol. 15, no. 10, pp. 2411–2423, Oct. 2016.

[15] V. Aggarwal, A. A. Mahimkar, H. Ma, Z. Zhang, S. Aeron, and W. Willinger, "Inferring smartphone service quality using tensor methods," in *Proc. 12th Int. Conf. Netw. Service Manage.*, Oct./Nov. 2016, pp. 263–267.

[16] P. Jain, P. Netrapalli, and S. Sanghavi, "Low-rank matrix completion using alternating minimization," in *Proc. 45th Annu. Symp. Theory Comput.*, Jun. 2013, pp. 665–674.

[17] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, pp. 717–772, Dec. 2009.

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

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

[20] M. Ashraphijuo, X. Wang, and J. Zhang, "Low-rank data completion with very low sampling rate using Newton's method," *IEEE Trans. Signal Process.*, vol. 67, no. 7, pp. 1849–1859, Apr. 2019.

[21] R. Vidal, "Subspace clustering," *IEEE Signal Process. Mag.*, vol. 28, no. 3, pp. 52–68, Mar. 2011. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.432.7808&rep=rep1&type=pdf

[22] L. Parsons, E. Haque, and H. Liu, "Subspace clustering for high dimensional data: A review," *ACM SIGKDD Explor. Newslett.*, vol. 6, no. 1, pp. 90–105, Jun. 2004.

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

[24] D. Pimentel, R. Nowak, and L. Balzano, "On the sample complexity of subspace clustering with missing data," in *Proc. IEEE Workshop Stat. Signal Process. (SSP)*, Jun./Jul. 2014, pp. 280–283.

[25] R. Tomioka, K. Hayashi, and H. Kashima, "Estimation of low-rank tensors via convex optimization," Oct. 2010, *arXiv:1010.0789*. [Online]. Available: https://arxiv.org/abs/1010.0789

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

[27] B. Romera-Paredes and M. Pontil, "A new convex relaxation for tensor completion," in *Proc. Adv. Neural Inf. Process. Syst.*, 2013, pp. 2967–2975.

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

[29] X.-Y. Liu, S. Aeron, V. Aggarwal, and X. Wang, "Low-tubal-rank tensor completion using alternating minimization," Oct. 2016, *arXiv:1610.01690*. [Online]. Available: https://arxiv.org/abs/1610.01690

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

[31] A. Krishnamurthy and A. Singh, "Low-rank matrix and tensor completion via adaptive sampling," in *Proc. Adv. Neural Inf. Process. Syst.*, 2013, pp. 836–844.

[32] D. Goldfarb and Z. Qin, "Robust low-rank tensor recovery: Models and algorithms," *SIAM J. Matrix Anal. Appl.*, vol. 35, no. 1, pp. 225–253, 2014.

[33] B. Recht, "A simpler approach to matrix completion," *J. Mach. Learn. Res.*, vol. 12, pp. 3413–3430, Dec. 2011.

[34] D. Pimentel-Alarcón, N. Boston, and R. D. Nowak, "A characterization of deterministic sampling patterns for low-rank matrix completion," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 4, pp. 623–636, Jun. 2016.

[35] M. Ashraphijuo, V. Aggarwal, and X. Wang, "On deterministic sampling patterns for robust low-rank matrix completion," *IEEE Signal Process. Lett.*, vol. 25, no. 3, pp. 343–347, Mar. 2018.

[36] F. J. Király, L. Theran, and R. Tomioka, "The algebraic combinatorial approach for low-rank matrix completion," *Mach. Learn.*, vol. 16, pp. 1391–1436, Jan. 2015.

[37] M. Ashraphijuo and X. Wang, "A characterization of sampling patterns for union of low-rank subspaces retrieval problem," in *Proc. Int. Symp. Artif. Intell. Math.*, 2018, pp. 1–8.

[38] L. Balzano, B. Eriksson, and R. Nowak, "High rank matrix completion and subspace clustering with missing data," in *Proc. Conf. Artif. Intell. Statist. (AIStats)*, 2012, pp. 1–15.

[39] M. Ashraphijuo, X. Wang, and V. Aggarwal, "Rank determination for low-rank data completion," *J. Mach. Learn. Res.*, vol. 18, no. 1, pp. 3422–3450, Jan. 2017.

[40] M. Ashraphijuo and X. Wang, "Clustering a union of low-rank subspaces of different dimensions with missing data," *Pattern Recognit. Lett.*, vol. 120, pp. 31–35, Apr. 2019.

[41] D. Pimentel-Alarcón and R. Nowak, "The information-theoretic requirements of subspace clustering with missing data," in *Proc. Int. Conf. Mach. Learn.*, Jun. 2016, pp. 802–810.

[42] C. Yang, D. Robinson, and R. Vidal, "Sparse subspace clustering with missing entries," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2015, pp. 2463–2472.

[43] M. Ashraphijuo, X. Wang, and V. Aggarwal, "An approximation of the CP-rank of a partially sampled tensor," in *Proc. Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2017, pp. 604–611.

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

[45] T. G. Kolda, "Orthogonal tensor decompositions," *SIAM J. Matrix Anal. Appl.*, vol. 23, no. 1, pp. 243–255, 2001.

[46] L. Grasedyck, "Hierarchical singular value decomposition of tensors," *SIAM J. Matrix Anal. Appl.*, vol. 31, no. 4, pp. 2029–2054, 2010.
[47] J. M. F. ten Berge and N. D. Sidiropoulos, "On uniqueness in candecomp/parafac," *Psychometrika*, vol. 67, no. 3, pp. 399–409, Sep. 2002.
[48] A. Stegeman and N. D. Sidiropoulos, "On Kruskal's uniqueness condition for the Candecomp/Parafac decomposition," *Linear Algebra Appl.*, vol. 420, no. 2, pp. 540–552, Jan. 2007.
[49] I. V. Oseledets, "Tensor-train decomposition," *SIAM J. Sci. Comput.*, vol. 33, no. 5, pp. 2295–2317, 2011.
[50] S. Holtz, T. Rohwedder, and R. Schneider, "On manifolds of tensors of fixed tt-rank," *Numerische Math.*, vol. 120, no. 4, pp. 701–731, Apr. 2012.
[51] J. Ballani, L. Grasedyck, and M. Kluge, "Black box approximation of tensors in hierarchical Tucker format," *Linear Algebra Appl.*, vol. 438, no. 2, pp. 639–657, Jan. 2013.
[52] W. Hackbusch and S. Kühn, "A new scheme for the tensor representation," *J. Fourier Anal. Appl.*, vol. 15, no. 5, pp. 706–722, Oct. 2009.
[53] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover, "Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 1, pp. 148–172, 2013.
[54] B. Sturmfels, *System of Polynomial Equations*. Rhode Island, United States Amer. Math. Soc., 2002.
[55] A. Schaeffer, "Inequalities of A. Markoff and S. Bernstein for polynomials and related functions," *Bull. Amer. Math. Soc.*, vol. 47, no. 8, pp. 565–579, 1941.
[56] M. Ashraphijuo and X. Wang, "Fundamental conditions for low-CP-rank tensor completion," *J. Mach. Learn. Res.*, vol. 18, no. 1, pp. 2116–2145, Jan. 2017.
[57] S. Janson, "On concentration of probability," *Contemp. Combinatorics*, vol. 10, no. 3, pp. 1–9, May 2002.
[58] H. A. Kierstead, "An effective version of Hall's theorem," *Proc. Amer. Math. Soc.*, vol. 88, no. 1, pp. 124–128, 1983.

Morteza Ashraphijuo received the B.Sc. degree in Electrical Engineering from Sharif University of Technology and M.Sc. degree in Industrial Engineering & Operations Research from University of California, Berkeley, in 2015 and 2016, respectively. He is currently pursuing the Ph.D. degree in Electrical Engineering from Columbia University in New York. His research interests lie in the general areas of machine learning, optimization theory and statistical learning. He has won multiple awards including a Silver Medal in the International Mathematical Olympiad (IMO) in 2010 and the Katta G. Murty Best Paper Prize from University of California, Berkeley, in 2016.

Vaneet Aggarwal (S'08–M'11–SM'15) received the B.Tech. degree in 2005 from the Indian Institute of Technology, Kanpur, India, and the M.A. and Ph.D. degrees in 2007 and 2010, respectively from Princeton University, Princeton, NJ, USA, all in Electrical Engineering.

He is currently an Associate Professor at Purdue University, West Lafayette, IN. Prior to this, he was a Senior Member of Technical Staff Research at AT&T Labs-Research, NJ (2010–2014). He was an Adjunct Assistant Professor at Columbia University, NY (2013–2014) and a VAJRA Adjunct Professor at IISc Bangalore (2018–2019). His research interests are in communications and networking, video streaming, cloud computing, and machine learning. He received Princeton University's Porter Ogden Jacobus Honorific Fellowship in 2009, AT&T Vice President Excellence Award in 2013, AT&T Senior Vice President Excellence Award in 2014, the 2017 Jack Neubauer Memorial Award recognizing the Best Systems Paper published in the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and the 2018 Infocom Workshop HotPOST Best Paper Award. He is serving on the editorial board of the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING.

Xiaodong Wang (S'98–M'98–SM'04–F'08) received the Ph.D. degree in Electrical Engineering from Princeton University. He is a Professor of Electrical Engineering at Columbia University in New York. Dr. Wang's research interests fall in the general areas of computing, signal processing and communications, and has published extensively in these areas. Among his publications is a book entitled *Wireless Communication Systems: Advanced Techniques for Signal Reception*, published by Prentice Hall in 2003. His current research interests include wireless communications, statistical signal processing, and genomic signal processing. Dr. Wang received the 1999 NSF CAREER Award, the 2001 IEEE Communications Society and Information Theory Society Joint Paper Award, and the 2011 IEEE Communication Society Award for Outstanding Paper on New Communication Topics. He has served as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the IEEE TRANSACTIONS ON INFORMATION THEORY. He is a Fellow of the IEEE and listed as an ISI Highly-cited Author.