Journal of Machine Learning Research 18 (2017) 1-29 Submitted 7/17; Revised 8/17; Published 9/17

Rank Determination for Low-Rank Data Completion

Morteza Ashraphijuo ASHRAPHIJUO@EE.COLUMBIA.EDU
Columbia University
New York, NY 10027, USA

Xiaodong Wang WANGX@EE.COLUMBIA.EDU
Columbia University
New York, NY 10027, USA

Vaneet Aggarwal VANEET@PURDUE.EDU
Purdue University
West Lafayette, IN 47907, USA

Editor: Animashree Anandkumar

Abstract

Recently, fundamental conditions on the sampling patterns have been obtained for finite completability of low-rank matrices or tensors given the corresponding ranks. In this paper, we consider the scenario where the rank is not given and we aim to approximate the unknown rank based on the location of sampled entries and some given completion. We consider a number of data models, including single-view matrix, multi-view matrix, CP tensor, tensor-train tensor and Tucker tensor. For each of these data models, we provide an upper bound on the rank when an arbitrary low-rank completion is given. We characterize these bounds both deterministically, i.e., with probability one given that the sampling pattern satisfies certain combinatorial properties, and probabilistically, i.e., with high probability given that the sampling probability is above some threshold. Moreover, for both single-view matrix and CP tensor, we are able to show that the obtained upper bound is exactly equal to the unknown rank if the lowest-rank completion is given. Furthermore, we provide numerical experiments for the case of single-view matrix, where we use nuclear norm minimization to find a low-rank completion of the sampled data and we observe that in most of the cases the proposed upper bound on the rank is equal to the true rank.

Keywords: Low-rank data completion, rank estimation, tensor, matrix, manifold, Tucker rank, tensor-train rank, CP rank, multi-view matrix.

1. Introduction

Developing methods and algorithms to study large high-dimensional data is becoming more indispensable as hyperspectral images and videos, product ranking datasets and other applications of big datasets are attracting more attention recently. Moreover, in order to guarantee the same level of efficiency in images or videos, a minor increment in dimensionality in the datasets entails a significant increment in the amount of the data, and this fact causes modeling and also computational challenges to analyze big high-dimensional datasets. Consequently, providing a statistically rigorous result requires a massive amount of data that grows exponentially with the dimension. The low-rank data completion problem is concerned with completing a matrix or tensor given a subset of its entries and some rank constraints. Various applications can be found in many fields including image and signal processing (Candès et al., 2013; Ji et al., 2010), data mining (Eldén, 2007), network coding (Harvey et al., 2005), compressed sensing (Lim and Comon, 2010; Sidiropoulos and Kyrillidis, 2012; Ashraphijuo et al., 2016c; Gandy et al., 2011; Ashraphijuo and Wang, 2017c; Ashraphijuo et al., 2015), reconstructing the visual data (Liu et al., 2013), bioinformatics and systems biology (Ogundijo et al., 2017; Ogundijo et al.), fingerprinting (Liu et al., 2016), etc. There is an extensive literature on developing various optimization methods to treat this problem including minimizing a convex relaxation of rank (Candès and Recht, 2009; Candès and Tao, 2010; Cai et al., 2010; Gandy et al., 2011; Ashraphijuo et al., 2016b; Ashraphijuo and Wang, 2017c), non-convex approaches (Recht and Ré, 2013), and alternating minimization (Jain et al., 2013; Ge et al., 2016), etc. More recently, deterministic conditions on the sampling patterns have been studied for subspace clustering in (Pimentel-Alarcón et al., 2016c, 2015, 2016a,b). Moreover, the fundamental conditions on the sampling pattern that lead to different numbers of completion (unique, finite, or infinite) for different data structures given the corresponding rank constraints have been investigated in (Pimentel-Alarcón et al., 2016d; Ashraphijuo et al., 2017c; Ashraphijuo and Wang, 2017b; Ashraphijuo et al., 2016a; Ashraphijuo and Wang, 2017a; Ashraphijuo et al., 2017d,a,b).

However, in many practical low-rank data completion problems, the rank may not be known *a priori*. In this paper, we investigate this problem and we aim to approximate the rank based on the given entries, where it is assumed that the original data is generically chosen from the manifold corresponding to the unknown rank. The only existing work that treats this problem for a single-view matrix data based on the sampling pattern is (Pimentel-Alarcón and Nowak, 2016), which requires some strong assumptions including the existence of a completion whose rank r is a lower bound on the unknown true rank r^*, i.e., $r^* \geq r$. We start by investigating the single-view matrix to provide a new analysis that does not require such assumption and also we can extend our approach to treat the CP rank tensor model. Moreover, we further generalize our approach to treat vector rank data models including the multi-view matrix, the Tucker rank tensor and the tensor-train (TT) rank tensor. For each of these data models, we obtain the upper bound on the scalar rank or component-wise upper bound on the unknown vector rank, deterministically based on the sampling pattern and the rank of a given completion. We also obtain such bound that holds with high probability based on the sampling probability. Moreover, for the single-view matrix, we provide some numerical results to show how tight our probabilistic bounds on the rank are (in terms of the sampling probability). In particular, we used nuclear norm minimization to find a completion and demonstrate our proposed method in obtaining a tight bound on the unknown rank.

In general, providing a completion requires much less samples than recovering the original sampled data. The goal of this paper is to solve the fundamental problem of rank determination for the original sampled data given an arbitrary low-rank data completion. One possible application scenario is to improve upon the low-rank completion obtained by the convex relaxation methods. Specifically, using convex optimization (minimization of nuclear and atomic norms or summation of nuclear norms of matricizations and unfoldings) or any other methods in the literature, we may be able to find a fairly low-rank "completion" of the original data, which is not necessarily equal (or even close) to the original sampled data. Then, under some circumstances, the rank of the obtained completion using any rank independent method can be an upper bound on the rank of the original sampled data (and sometimes the obtained rank can be exactly equal to the rank of the original sampled data).

We take advantage of the geometric analysis on the manifold of the corresponding data which leads to the fundamental conditions on the sampling pattern (independent of the value of entries) (Pimentel-Alarcón et al., 2016d; Ashraphijuo et al., 2017c; Ashraphijuo and Wang, 2017b; Ashraphijuo et al., 2016a; Ashraphijuo and Wang, 2017a) such that given an arbitrary low-rank completion we can provide a tight upper bound on the rank. To illustrate how such approximation is even possible consider the following example. Assume that an $n_1 \times n_2$ rank-2 matrix is chosen generically from the corresponding manifold. Hence, any 2×2 submatrix of this matrix is full-rank with probability one (due to the genericity assumption). Moreover, note that any 3×3 submatrix of this matrix is not full-rank. As a result, by observing the sampled entries we can find some bounds on the rank. Using the analysis in (Pimentel-Alarcón et al., 2016d; Ashraphijuo et al., 2017c; Ashraphijuo and Wang, 2017b; Ashraphijuo et al., 2016a; Ashraphijuo and Wang, 2017a) on finite completablity of the sampled data (finite number of completions) for different data models, we characterize both deterministic and probablistic bounds on the unknown rank.

The remained of the paper is organized as follows. In Section 2, we introduce the data models and problem statement. In Sections 3 and 4 we characterize our determintic and probablistic bounds for scalar-rank cases (single-view matrix and CP tensor) and vector-rank cases (multi-view matrix, Tucker tensor and TT tensor), respectively. Finally, Section 5 concludes the paper.

2. Data Models and Problem Statement

2.1 Matrix Models

2.1.1 SINGLE-VIEW MATRIX

Assume that the sampled matrix $\mathbf{U}$ is chosen generically from the manifold of the $n_1 \times n_2$ matrices of rank r^*, where r^* is unknown. The matrix $\mathbf{V} \in \mathbb{R}^{n_1 \times r^*}$ is called a basis for $\mathbf{U}$ if each column of $\mathbf{U}$ can be written as a linear combination of the columns of $\mathbf{V}$. Denote $\mathbf{\Omega}$ as the binary sampling pattern matrix that is of the same size as $\mathbf{U}$ and $\mathbf{\Omega}(\vec{x}) = 1$ if $\mathbf{U}(\vec{x})$ is observed and $\mathbf{\Omega}(\vec{x}) = 0$ otherwise, where $\vec{x} = (x_1, x_2)$ represents the entry corresponding to row number x_1 and column number x_2. Moreover, define $\mathbf{U_\Omega}$ as the matrix obtained from sampling $\mathbf{U}$ according to $\mathbf{\Omega}$, i.e.,

$$\mathbf{U_\Omega}(\vec{x}) = \begin{cases} \mathbf{U}(\vec{x}) & \text{if } \mathbf{\Omega}(\vec{x}) = 1, \\ 0 & \text{if } \mathbf{\Omega}(\vec{x}) = 0. \end{cases} \tag{1}$$

2.1.2 MULTI-VIEW MATRIX

The matrix $\mathbf{U} \in \mathbb{R}^{n \times (n_1+n_2)}$ is sampled. Denote a partition of $\mathbf{U}$ as $\mathbf{U} = [\mathbf{U}_1|\mathbf{U}_2]$ where $\mathbf{U}_1 \in \mathbb{R}^{n \times n_1}$ and $\mathbf{U}_2 \in \mathbb{R}^{n \times n_2}$ represent the first and second views of data, respectively. The sampling pattern is defined as $\mathbf{\Omega} = [\mathbf{\Omega}_1|\mathbf{\Omega}_2]$, where $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$ represent the sampling patterns corresponding to the first and second views of data, respectively. Assume that $\text{rank}(\mathbf{U}_1) = r_1^*$, $\text{rank}(\mathbf{U}_2) = r_2^*$ and $\text{rank}(\mathbf{U}) = r^*$, and also $\mathbf{U}$ is chosen generically from the manifold structure with above parameters. Denote $\underline{r}^* = (r_1^*, r_2^*, r^*)$ which is assumed unknown.

2.2 Tensor Models

Assume that a d-way tensor $\mathcal{U} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ is sampled. For the sake of simplicity in notation, define $N_i \triangleq \left(\Pi_{j=1}^{i} n_j\right)$, $\bar{N}_i \triangleq \left(\Pi_{j=i+1}^{d} n_j\right)$ and $N_{-i} \triangleq \frac{N_d}{n_i}$. 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, where $\mathcal{U}(\vec{x})$ represents an entry of tensor $\mathcal{U}$ with coordinate $\vec{x} = (x_1, \ldots, x_d)$. Moreover, define $\mathcal{U}_\Omega$

as the tensor obtained from sampling $\mathcal{U}$ according to Ω, i.e.,

$$\mathcal{U}_\Omega(\vec{x}) = \begin{cases} \mathcal{U}(\vec{x}) & \text{if } \Omega(\vec{x}) = 1, \\ 0 & \text{if } \Omega(\vec{x}) = 0. \end{cases} \tag{2}$$

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.

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 $\mathcal{U}(\vec{x}) = \widetilde{\mathbf{U}}_{(i)}(\widetilde{M}_i(x_1,\dots,x_i), \widetilde{M}_{-i}(x_{i+1},\dots,x_d))$, where $\widetilde{M}_i : (x_1,\dots,x_i) \rightarrow \{1,2,\dots,N_i\}$ and $\widetilde{M}_{-i} : (x_{i+1},\dots,x_d) \rightarrow \{1,2,\dots,\bar{N}_i\}$ are two bijective mappings.

Let $\mathbf{U}_{(i)} \in \mathbb{R}^{n_i \times N_{-i}}$ be the i-th *matricization* of the tensor $\mathcal{U}$, such that $\mathcal{U}(\vec{x}) = \mathbf{U}_{(i)}(x_i, M_i(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d))$, where $M_i : (x_1,\dots,x_{i-1},x_{i+1},\dots,x_d) \rightarrow \{1,2,\dots,N_{-i}\}$ is a bijective mapping. Observe that for any arbitrary tensor $\mathcal{A}$, the first matricization and the first unfolding are the same, i.e., $\mathbf{A}_{(1)} = \widetilde{\mathbf{A}}_{(1)}$.

In what follows, we introduce three different tensor ranks, i.e., the CP rank, Tucker rank and TT rank.

2.2.1 CP DECOMPOSITION

The CP rank of a tensor $\mathcal{U}$, $\text{rank}_{\text{CP}}(\mathcal{U}) = r$, is defined as the minimum number r such that there exist $\mathbf{a}_i^l \in \mathbb{R}^{n_i}$ for $1 \leq i \leq d$ and $1 \leq l \leq r$, such that

$$\mathcal{U} = \sum_{l=1}^{r} \mathbf{a}_1^l \otimes \mathbf{a}_2^l \otimes \ldots \otimes \mathbf{a}_d^l, \tag{3}$$

or equivalently,

$$\mathcal{U}(x_1, x_2, \dots, x_d) = \sum_{l=1}^{r} \mathbf{a}_1^l(x_1)\mathbf{a}_2^l(x_2)\ldots\mathbf{a}_d^l(x_d), \tag{4}$$

where $\otimes$ denotes the tensor product (outer product) and $\mathbf{a}_i^l(x_i)$ denotes the x_i-th entry of vector $\mathbf{a}_i^l$. Note that $\mathbf{a}_1^l \otimes \mathbf{a}_2^l \otimes \ldots \otimes \mathbf{a}_d^l \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ is a rank-1 tensor, $l = 1, 2, \dots, r$.

2.2.2 TUCKER DECOMPOSITION

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

$$\mathcal{U}'(x_1, \cdots, x_{i-1}, k_i, x_{i+1}, \cdots, x_d) \triangleq \sum_{x_i=1}^{n_i} \mathcal{U}(x_1, \cdots, x_{i-1}, x_i, x_{i+1}, \cdots, x_d)\mathbf{X}(x_i, k_i). \tag{5}$$

The Tucker rank of a tensor $\mathcal{U}$ is defined as $\text{rank}_{\text{Tucker}}(\mathcal{U}) = \underline{r} = (m_1, \dots, m_d)$ where $m_i = \text{rank}(\mathbf{U}_{(i)})$, i.e., the rank of the i-th matricization, $i = 1, \dots, d$. The Tucker decomposition of $\mathcal{U}$ is given by

$$\mathcal{U}(\vec{x}) = \sum_{k_1=1}^{m_1} \cdots \sum_{k_d=1}^{m_d} \mathcal{C}(k_1, \dots, k_d)\mathbf{T}_1(k_1, x_1) \ldots \mathbf{T}_d(k_d, x_d), \tag{6}$$

or in short

$$\mathcal{U} = \mathcal{C} \times_{i=1}^{d} \mathbf{T}_i, \tag{7}$$

where $\mathcal{C} \in \mathbb{R}^{m_1 \times \cdots \times m_d}$ is the core tensor and $\mathbf{T}_i \in \mathbb{R}^{m_i \times n_i}$ are d orthogonal matrices.

2.2.3 TT DECOMPOSITION

The separation or TT rank of a tensor is defined as $\text{rank}_{\text{TT}}(\mathcal{U}) = \underline{r} = (u_1, \ldots, u_{d-1})$ where $u_i = \text{rank}(\widetilde{\mathbf{U}}_{(i)})$, i.e., the rank of the i-th unfolding, $i = 1, \ldots, d-1$. Note that $u_i \leq \max\{N_i, \bar{N}_i\}$ in general and also u_1 is simply the conventional matrix rank when $d = 2$. The TT decomposition of a tensor $\mathcal{U}$ is given by

$$\mathcal{U}(\vec{x}) = \sum_{k_1=1}^{u_1} \cdots \sum_{k_{d-1}=1}^{u_{d-1}} \mathcal{U}^{(1)}(x_1, k_1) \left(\prod_{i=2}^{d-1} \mathcal{U}^{(i)}(k_{i-1}, x_i, k_i) \right) \mathcal{U}^{(d)}(k_{d-1}, x_d), \tag{8}$$

or in short

$$\mathcal{U} = \mathcal{U}^{(1)} \ldots \mathcal{U}^{(d)}, \tag{9}$$

where the 3-way tensors $\mathcal{U}^{(i)} \in \mathbb{R}^{u_{i-1} \times n_i \times u_i}$ for $i = 2, \ldots, d-1$ and matrices $\mathcal{U}^{(1)} \in \mathbb{R}^{n_1 \times u_1}$ and $\mathcal{U}^{(d)} \in \mathbb{R}^{u_{d-1} \times n_d}$ are the components of this decomposition.

For each matrix or tensor model, we assume that the true rank of $\mathbf{U}$ or $\mathcal{U}$ is r^* or $\underline{r}^*$ which is unknown, and also $\mathbf{U}$ or $\mathcal{U}$ is chosen generically from the corresponding manifold. The table below represents the mentioned symbols in brief.

Data structure	Sampled data	Rank	Comments
Single-view matrix	$\mathbf{U} \in \mathbb{R}^{n_1 \times n_2}$	r^*	–
Multi-view matrix	$\mathbf{U} = [\mathbf{U}_1 \vert \mathbf{U}_2] \in \mathbb{R}^{n \times (n_1+n_2)}$	$\underline{r}^* = (r_1^*, r_2^*, r^*)$	$r_i^* = \text{rank}(\mathbf{U}_i)$
CP tensor	$\mathcal{U} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$	r^*	–
Tucker tensor	$\mathcal{U} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$	$\underline{r}^* = (m_1^*, \ldots, m_d^*)$	$m_i^* = \text{rank}(\mathbf{U}_{(i)})$
TT tensor	$\mathcal{U} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$	$\underline{r}^* = (u_1^*, \ldots, u_{d-1}^*)$	$u_i^* = \text{rank}(\widetilde{\mathbf{U}}_{(i)})$

2.3 Problem Statement

In this paper, we assume that there exists a full rank completion of the sampled data (i.e., the data is not over-sampled). For each one of the above data models, we are interested in obtaining the upper bound on the unknown scalar-rank r^* or component-wise upper bound on the unknown vector-rank $\underline{r}^*$, deterministically based on the sampling pattern $\mathbf{\Omega}$ or Ω and the rank of a given completion. Also, we aim to provide such bound that holds with high probability based only on the sampling probability of the entries and the rank of a given completion. Moreover, for the single-view matrix model and CP-rank tensor model, where the rank is a scalar, we provide both deterministic and probabilistic conditions such that the unknown rank can be exactly determined.

3. Scalar-Rank Cases

3.1 Single-View Matrix

Previously, this problem has been treated in (Pimentel-Alarcón and Nowak, 2016), where strong assumptions including the existence of a completion with rank $r \leq r^*$ have been used. In this section, we provide an analysis that does not require such assumption and moreover our analysis can be extended to multi-view data and tensors in the following sections. Furthermore, we show the tightness of our theoretical bounds via numerical examples.

Assume that $\mathbf{U} \in \mathbb{R}^{n_1 \times n_2}$ is the sampled matrix. Let $\mathbb{P}_1$ and $\mathbb{P}_2$ denote the Lebesgue measures on $\mathbb{R}^{n_1 \times r^*}$ and $\mathbb{R}^{r^* \times n_2}$, respectively. In this paper, we assume that the matrix $\mathbf{U}$ is chosen generically from the manifold of $n_1 \times n_2$ matrices of rank r^*, i.e., the entries of $\mathbf{U}$ are drawn independently with respect to Lebesgue measure on the corresponding manifold. Hence, the probability measures of all statements in this subsection are $\mathbb{P}_1 \times \mathbb{P}_2$.

3.1.1 Deterministic Rank Analysis

The following condition will be used frequently in this subsection.

Condition A_r: Each column of the sampled matrix includes at least r sampled entries.

Consider an arbitrary column of the sampled matrix $\mathbf{U}(:, i)$, where $i \in \{1, \dots, n_2\}$. Let $l_i = N_{\mathbf{\Omega}}(\mathbf{U}(:, i))$ denote the number of observed entries in the i-th column of $\mathbf{U}$. Condition A_r results that $l_i \geq r$.

We construct a binary valued matrix called **constraint matrix** $\breve{\mathbf{\Omega}}_r$ based on $\mathbf{\Omega}$ and a given number r. Specifically, we construct $l_i - r$ columns with binary entries based on the locations of the observed entries in $\mathbf{U}(:, i)$ such that each column has exactly $r+1$ entries equal to one. Assume that $x_1, \dots, x_{l_i}$ are the row indices of all observed entries in this column. Let $\mathbf{\Omega}_r^i$ be the corresponding $n_1 \times (l_i - r)$ matrix to this column which is defined as the following: for any $j \in \{1, \dots, l_i - r\}$, the j-th column has the value 1 in rows $\{x_1, \dots, x_r, x_{r+j}\}$ and zeros elsewhere. Define the binary constraint matrix as $\breve{\mathbf{\Omega}}_r = \left[\mathbf{\Omega}_r^1 | \mathbf{\Omega}_r^2 \dots | \mathbf{\Omega}_r^{n_2}\right] \in \mathbb{R}^{n_1 \times K_r}$ (Pimentel-Alarcón et al., 2016d), where $K_r = N_{\mathbf{\Omega}}(\mathbf{U}) - n_2 r$.

Condition B_r: There exists a submatrix[1] $\breve{\mathbf{\Omega}}'_r \in \mathbb{R}^{n_1 \times K}$ of $\breve{\mathbf{\Omega}}_r$ such that $K = n_1 r - r^2$ and for any $K' \in \{1, 2, \dots, K\}$ and any submatrix $\breve{\mathbf{\Omega}}''_r \in \mathbb{R}^{n_1 \times K'}$ of $\breve{\mathbf{\Omega}}'_r$ we have

$$r f(\breve{\mathbf{\Omega}}''_r) - r^2 \geq K', \tag{10}$$

where $f(\breve{\mathbf{\Omega}}''_r)$ denotes the number of nonzero rows of $\breve{\mathbf{\Omega}}''_r$.

Note that exhaustive enumeration is needed in order to check whether or not Condition B_r holds. Hence, the deterministic analysis cannot be used in practice for large-scale data. However, it serves as the basis of the subsequent probabilistic analysis that will lead to a simple lower bound on the sampling probability such that Condition B_r holds with high probability, which is of practical value.

In the following, we restate Theorem 1 in (Pimentel-Alarcón et al., 2016d) which will be used later.

Lemma 1 *With probability one, there are finitely many completions of the sampled matrix if and only if Conditions A_{r^*} and B_{r^*} hold.*

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

Recall that the true rank r^* is assumed unknown.

Definition 2 *Let $\mathcal{S}_{\mathbf{\Omega}}$ denote the set of all natural numbers r such that both Conditions A_r and B_r hold.*

Lemma 3 *There exists a number $r_{\mathbf{\Omega}}$ such that $\mathcal{S}_{\mathbf{\Omega}} = \{1, 2, \ldots, r_{\mathbf{\Omega}}\}$.*

Proof Assume that $1 < r \leq \min\{n_1, n_2\}$ and $r \in \mathcal{S}_{\mathbf{\Omega}}$. It suffices to show $r - 1 \in \mathcal{S}_{\mathbf{\Omega}}$. By contradiction, assume that $r - 1 \notin \mathcal{S}_{\mathbf{\Omega}}$. Therefore, according to Lemma 1, there exist infinitely many completions of $\mathbf{U}$ of rank $r - 1$ and there exist at most finitely many completions of $\mathbf{U}$ of rank r.

Consider a rank $r - 1$ completion $\mathbf{U}_{r-1}$. Note that changing one single entry (a non-observed entry) of $\mathbf{U}_{r-1}$, for example $\mathbf{U}_{r-1}(1, 1) = x$, to a random number in $y \in \mathbb{R}$ changes the rank of this matrix by at most 1 and basically since we are changing to a random number, it can be easily seen that the rank does not decrease with probability one. Hence, the rank of the modified matrix $\mathbf{U}'_{r-1}$ would be either $r - 1$ or r. Assume that the rank has been increased to r. Then, we show there exist infinitely many completions of rank r, which contradicts the assumption. In fact, for any value of $\mathbf{U}_{r-1}(1, 1)$ except x, this matrix would be a rank r completion. To observe this more clearly, consider the $r \times r$ submatrix of $\mathbf{U}'_{r-1}$ whose determinant is not zero due to changing the value of $\mathbf{U}_{r-1}(1, 1)$. It is easily observed that this submatrix includes $\mathbf{U}'_{r-1}(1, 1)$ and let assume it is $\mathbf{U}'_{r-1}(1 : r, 1 : r)$, and therefore the determinant of $\mathbf{U}'_{r-1}(2 : r, 2 : r)$ is a nonzero number (otherwise the rank would not increase by changing the value of $\mathbf{U}_{r-1}(1, 1)$). Hence, it is easy to see that for any value of $\mathbf{U}_{r-1}(1, 1)$ except x, $\mathbf{U}'_{r-1}$ would be a rank r completion, and therefore there exist infinitely many completions of rank r and proof is complete in this scenario.

Now, assume that changing any of the non-observed entries does not increase the rank of $\mathbf{U}_{r-1}$. Then, this contradicts the assumption that there exists a full rank completion of the sampled data since there does not exist any completion of rank higher than $r - 1$. ■

The following theorem provides a relationship between the unknown rank r^* and $r_{\mathbf{\Omega}}$.

Theorem 4 *With probability one, exactly one of the following statements holds*

(i) $r^* \in \mathcal{S}_{\mathbf{\Omega}} = \{1, 2, \ldots, r_{\mathbf{\Omega}}\}$;

(ii) For any arbitrary completion of the sampled matrix $\mathbf{U}$ of rank r, we have $r \notin \mathcal{S}_{\mathbf{\Omega}}$.

Proof Suppose that there does not exist a completion of the sampled matrix $\mathbf{U}$ of rank r such that $r \in \mathcal{S}_{\mathbf{\Omega}}$. Therefore, it is easily verified that statement (ii) holds and statement (i) does not hold. On the other hand, assume that there exists a completion of the sampled matrix $\mathbf{U}$ of rank r, where $r \in \mathcal{S}_{\mathbf{\Omega}}$. Hence, statement (ii) does not hold and to complete the proof it suffices to show that with probability one, statement (i) holds.

Observe that $r_{\mathbf{\Omega}} \in \mathcal{S}_{\mathbf{\Omega}}$, and therefore Condition $A_{r_{\mathbf{\Omega}}}$ holds. Hence, each column of $\mathbf{U}$ includes at least $r_{\mathbf{\Omega}} + 1$ observed entries. On the other hand, the existence of a completion of the sampled matrix $\mathbf{U}$ of rank $r \in \mathcal{S}_{\mathbf{\Omega}}$ results in the existence of a basis $\mathbf{X} \in \mathbb{R}^{n_1 \times r}$ such that each column of $\mathbf{U}$ is a linear combination of the columns of $\mathbf{X}$, and thus there exists $\mathbf{Y} \in \mathbb{R}^{r \times n_2}$ such that $\mathbf{U}_{\mathbf{\Omega}} = (\mathbf{X}\mathbf{Y})_{\mathbf{\Omega}}$. Hence, given $\mathbf{X}$, each observed entry $\mathbf{U}(i, j)$ results in a degree-1 polynomial in terms of the entries of $\mathbf{Y}$ as the following

$$\mathbf{U}(i,j) = \sum_{l=1}^{r} \mathbf{X}(i,l)\mathbf{Y}(l,j). \tag{11}$$

Consider the first column of $\mathbf{U}$ and recall that it includes at least $r_{\mathbf{\Omega}}+1 \geq r+1$ observed entries. The genericity of the coefficients of the above-mentioned polynomials results that using r of the observed entries the first column of $\mathbf{Y}$ can be determined uniquely. This is because there exists a unique solution for a system of r linear equations in r variables that are linearly independent. Then, there exists at least one more observed entry besides these r observed entries in the first column of $\mathbf{U}$ and it can be written as a linear combination of the r observed entries that have been used to obtain the first column of $\mathbf{Y}$. Let $\mathbf{U}(i_1,1)$, $\dots$, $\mathbf{U}(i_r,1)$ denote the r observed entries that have been used to obtain the first column of $\mathbf{Y}$ and $\mathbf{U}(i_{r+1},1)$ denote the other observed entry. Hence, the existence of a completion of the sampled matrix $\mathbf{U}$ of rank $r \in \mathcal{S}_{\mathbf{\Omega}}$ results in an equation as the following

$$\mathbf{U}(i_{r+1},1) = \sum_{l=1}^{r} t_l \mathbf{U}(i_l,1), \tag{12}$$

where t_l's are constant scalars, $l = 1,\dots,r$. Assume that $r^* \notin \mathcal{S}_{\mathbf{\Omega}}$, i.e., statement (i) does not hold. Then, note that $r^* \geq r+1$ and $\mathbf{U}$ is chosen generically from the manifold of $n_1 \times n_2$ rank-r^* matrices, and therefore an equation of the form of (12) holds with probability zero. Moreover, according to Lemma 1 there exist at most finitely many completions of the sampled matrix of rank r. Therefore, there exist a completion of $\mathbf{U}$ of rank r with probability zero, which contradicts the initial assumption that there exists a completion of the sampled matrix $\mathbf{U}$ of rank r, where $r \in \mathcal{S}_{\mathbf{\Omega}}$. ■

Note that the existing work that treats the similar problem for a single-view matrix data based on the sampling pattern is (Pimentel-Alarcón and Nowak, 2016), which requires some strong assumptions including the existence of a completion whose rank r is a lower bound on the unknown true rank r^*, i.e., $r^* \geq r$. We provide a new analysis that does not require such assumption and also based on our new analysis, we can extend our approach to treat other data structures.

Corollary 5 *Consider an arbitrary number $r' \in \mathcal{S}_{\mathbf{\Omega}}$. Similar to Theorem 4, it follows that with probability one, exactly one of the followings holds*
(i) $r^* \in \{1,2,\dots,r'\}$;
(ii) For any arbitrary completion of the sampled matrix $\mathbf{U}$ of rank r, we have $r \notin \{1,2,\dots,r'\}$.

As a result of Corollary 5, we have the following.

Corollary 6 *Assuming that there exists a rank-r completion of the sampled matrix $\mathbf{U}$ such that $r \in \mathcal{S}_{\mathbf{\Omega}}$, then with probability one $r^* \leq r$.*

Corollary 7 *Let $\mathbf{U}^*$ denote an optimal solution to the following NP-hard optimization problem*

$$\begin{aligned} &\text{minimize}_{\mathbf{U}' \in \mathbb{R}^{n_1 \times n_2}} && \text{rank}(\mathbf{U}') \\ &\text{subject to} && \mathbf{U}'_{\mathbf{\Omega}} = \mathbf{U}_{\mathbf{\Omega}}. \end{aligned} \tag{13}$$

Also, let $\hat{\mathbf{U}}$ denote a suboptimal solution to the above optimization problem. Then, Corollary 5 results the following statements:

(i) If $\text{rank}(\mathbf{U}^*) \in \mathcal{S}_{\mathbf{\Omega}}$, then $r^* = \text{rank}(\mathbf{U}^*)$ with probability one.

(ii) If $\text{rank}(\hat{\mathbf{U}}) \in \mathcal{S}_{\mathbf{\Omega}}$, then $r^* \leq \text{rank}(\hat{\mathbf{U}})$ with probability one.

Remark 8 *One challenge of applying Corollary 7 or any of the other obtained deterministic results is the computation of $\mathcal{S}_{\mathbf{\Omega}}$, which involves exhaustive enumeration to check Condition B_r. Next, for each number r, we provide a lower bound on the sampling probability in terms of r that ensures $r \in \mathcal{S}_{\mathbf{\Omega}}$ with high probability. Consequently, we do not need to compute $\mathcal{S}_{\mathbf{\Omega}}$ but instead we can certify the above results with high probability.*

3.1.2 PROBABILISTIC RANK ANALYSIS

The following lemma is a re-statement of Theorem 3 in (Pimentel-Alarcón et al., 2016d), which is the probabilistic version of Lemma 1.

Lemma 9 *Suppose $r \leq \frac{n_1}{6}$ and that each* **column** *of the sampled matrix is observed in at least l entries, uniformly at random and independently across entries, where*

$$l > \max\left\{12 \ \log\left(\frac{n_1}{\epsilon}\right) + 12, 2r\right\}. \tag{14}$$

Also, assume that $r(n_1 - r) \leq n_2$. Then, with probability at least $1 - \epsilon$, $r \in \mathcal{S}_{\mathbf{\Omega}}$.

The following lemma is taken from (Ashraphijuo et al., 2016a) and will be used to derive a lower bound on the sampling probability that leads to the similar statement as Theorem 4 with high probability.

Lemma 10 *Consider a vector with n entries where each entry is observed with probability p independently from the other entries. If $p > p' = \frac{k}{n} + \frac{1}{\sqrt[4]{n}}$, then with probability at least $\left(1 - \exp(-\frac{\sqrt{n}}{2})\right)$, more than k entries are observed.*

The following proposition characterizes the probabilistic version of Theorem 4.

Proposition 11 *Suppose $r \leq \frac{n_1}{6}$, $r(n_1 - r) \leq n_2$ and that each entry of the sampled matrix is observed uniformly at random and independently across entries with probability p, where*

$$p > \frac{1}{n_1} \max\left\{12 \ \log\left(\frac{n_1}{\epsilon}\right) + 12, 2r\right\} + \frac{1}{\sqrt[4]{n_1}}. \tag{15}$$

Then, with probability at least $(1 - \epsilon)\left(1 - \exp(-\frac{\sqrt{n_1}}{2})\right)^{n_2}$, we have $r \in \mathcal{S}_{\mathbf{\Omega}}$.

Proof Consider an arbitrary column of $\mathbf{U}$ and note that resulting from Lemma 10 the number of observed entries at this column of $\mathbf{U}$ is greater than $\max\left\{12 \ \log\left(\frac{n_1}{\epsilon}\right) + 12, 2r\right\}$ with probability at least $\left(1 - \exp(-\frac{\sqrt{n_1}}{2})\right)$. Therefore, the number of sampled entries at each column satisfies

$$l > \max\left\{12 \ \log\left(\frac{n_1}{\epsilon}\right) + 12, 2r\right\}, \tag{16}$$

with probability at least $\left(1 - \exp(-\frac{\sqrt{n_1}}{2})\right)^{n_2}$. Thus, resulting from Lemma 9 with probability at least
$(1-\epsilon)\left(1 - \exp(-\frac{\sqrt{n_1}}{2})\right)^{n_2}$, we have $r \in \mathcal{S}_{\mathbf{\Omega}}$. ■

Finally, we have the following probabilistic version of Corollary 7.

Corollary 12 *Assume that* $\text{rank}(\mathbf{U}^*) \leq \frac{n_1}{6}$ *and* $\text{rank}(\mathbf{U}^*)(n_1 - \text{rank}(\mathbf{U}^*)) \leq n_2$ *and* (15) *holds for* $r = \text{rank}(\mathbf{U}^*)$*, where* $\mathbf{U}^*$ *denotes an optimal solution to the optimization problem* (13)*. Then, according to Proposition 11 and Corollary 7, with probability at least* $(1-\epsilon)\left(1 - \exp(-\frac{\sqrt{n_1}}{2})\right)^{n_2}$, $r^* = \text{rank}(\mathbf{U}^*)$*. Similarly, assume that* $\text{rank}(\hat{\mathbf{U}}) \leq \frac{n_1}{6}$ *and* $\text{rank}(\hat{\mathbf{U}})(n_1 - \text{rank}(\hat{\mathbf{U}})) \leq n_2$ *and* (15) *holds for* $r = \text{rank}(\hat{\mathbf{U}})$*, where* $\hat{\mathbf{U}}$ *denotes a suboptimal solution to the optimization problem* (13)*. Then, with probability at least* $(1-\epsilon)\left(1 - \exp(-\frac{\sqrt{n_1}}{2})\right)^{n_2}$, $r^* \leq \text{rank}(\hat{\mathbf{U}})$.

3.1.3 NUMERICAL RESULTS

In Fig. 1 and Fig. 2, the x-axis represents the sampling probability, and the y-axis denotes the value of r. The color scale represents the lower bound on the probability of event $r \in \mathcal{S}_{\mathbf{\Omega}}$. For example, as we can observe in Fig. 1, for any $r \in \{1, \ldots, 44\}$ we have $r \in \mathcal{S}_{\mathbf{\Omega}}$ with probability at least 0.6 (approximately based on the color scale since the corresponding points are orange) given that $p = 0.54$.

We consider the sampled matrix $\mathbf{U} \in \mathbb{R}^{300 \times 15000}$ and $\mathbf{U} \in \mathbb{R}^{1200 \times 240000}$ in Fig. 1 and Fig. 2, respectively. In particular, for fixed values of sampling probability p and r, we first find a "small" ϵ that (15) holds by trial-and-error. Then, according to Proposition 11, we conclude that with probability at least $(1-\epsilon)\left(1 - \exp(-\frac{\sqrt{n_1}}{2})\right)^{n_2}$, $r \in \mathcal{S}_{\mathbf{\Omega}}$.

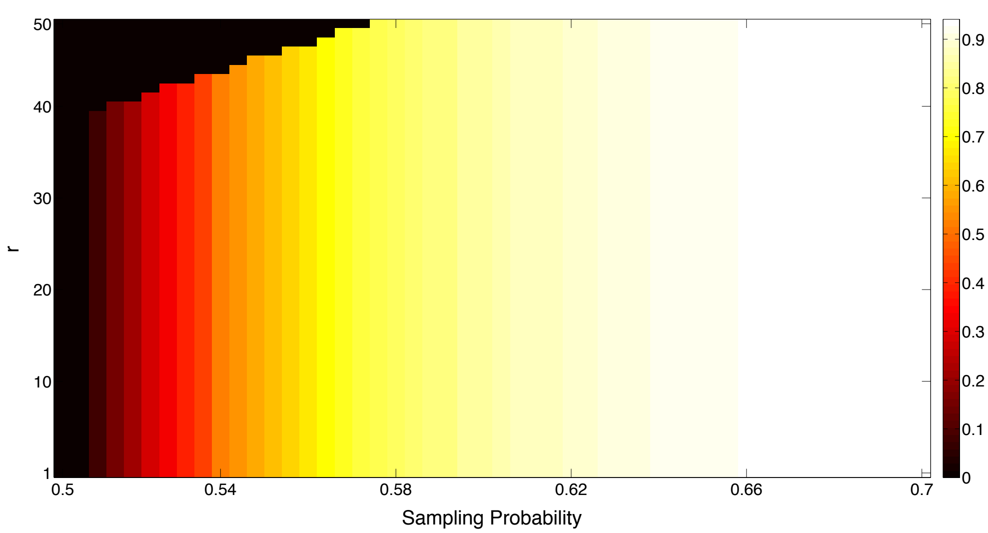

Figure 1: Probability of $r \in \mathcal{S}_{\mathbf{\Omega}}$ as a function of sampling probability for $\mathbf{U} \in \mathbb{R}^{300 \times 15000}$.

The purpose of Figs. 3–6 is to show how tight our proposed upper bounds on rank can be. Here, we first generate an $n_1 \times n_2$ random matrix of a given rank r by multiplying a random

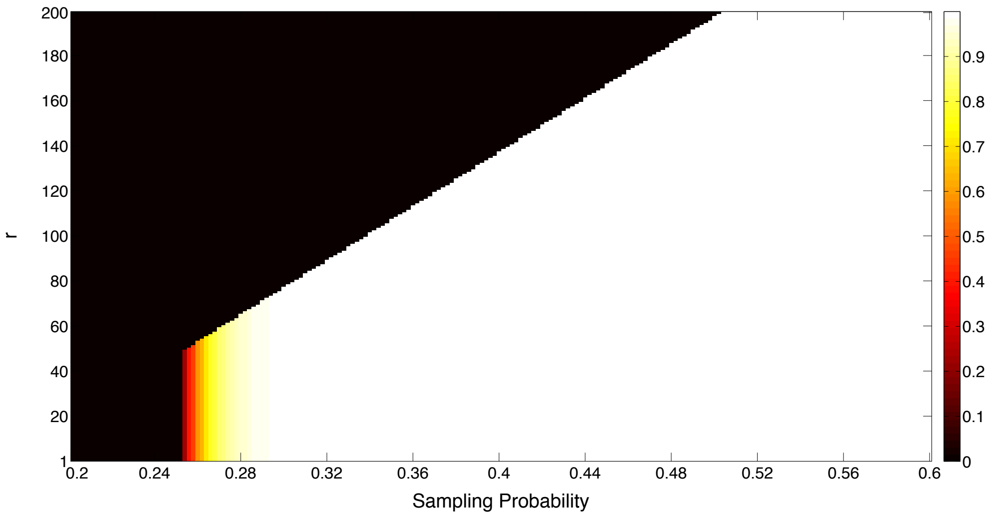

Figure 2: Probability of $r \in \mathcal{S}_{\mathbf{\Omega}}$ as a function of sampling probability for $\mathbf{U} \in \mathbb{R}^{1200\times 240000}$.

(entries are drawn according to a uniform distribution on real numbers within an interval) $n_1 \times r$ matrix and $r \times n_2$ matrix. Then, each entry of the randomly generated matrix is sampled uniformly at random and independently across entries with some sampling probability p. Afterwards, we apply the nuclear norm minimization method proposed in (Candes and Recht, 2012) for matrix completion, where the non-convex objective function in (13) is relaxed by using nuclear norm, which is the convex hull of the rank function, as follows

$$\begin{aligned} &\text{minimize}_{\mathbf{U}'\in\mathbb{R}^{n_1\times n_2}} && \|\mathbf{U}'\|_* \\ &\text{subject to} && \mathbf{U}'_{\mathbf{\Omega}} = \mathbf{U}_{\mathbf{\Omega}}, \end{aligned} \tag{17}$$

where $\|\mathbf{U}'\|_*$ denotes the nuclear norm of $\mathbf{U}'$. Let $\hat{\mathbf{U}}^*$ denote an optimal solution to (17) and recall that $\mathbf{U}^*$ denotes an optimal solution to (13). Since (17) is a convex relaxation to (13), we conclude that $\hat{\mathbf{U}}^*$ is a suboptimal solution to (13), and therefore $\text{rank}(\mathbf{U}^*) \leq \text{rank}(\hat{\mathbf{U}}^*)$. We used the Matlab program found online (Shabat, 2015) to solve (17).

As an example, we generate a random matrix $\mathbf{U} \in \mathbb{R}^{300\times 15000}$ (the same size as the matrix in Fig. 1) of rank r as described above for $r \in \{1,\dots,50\}$ and some values of the sampling probability p. Then, we obtain the rank of the completion given by (17) and denote it by r'. Due to the randomness of the sampled matrix, we repeat this procedure 5 times. We calculate the "gap" $r' - r$ in each of these 5 runs and denote the maximum and minimum among these 5 numbers by d_{max} and d_{min}, respectively. Hence, d_{max} and d_{min} represent the loosest (worst) and tightest (best) gaps between the rank obtained by (17) and rank of the original sampled matrix over 5 runs, respectively. In Figs. 3–6, the maximum and minimum gaps are plotted as a function of rank of the matrix, for different sampling probabilities.

We have the following observations.

- According to Fig. 1, for $p = 0.54$ and $p = 0.58$ we can ensure that the rank of any completion is an upper bound on the rank of the sampled matrix or r^* with probability at least 0.6 and 0.8, respectively.

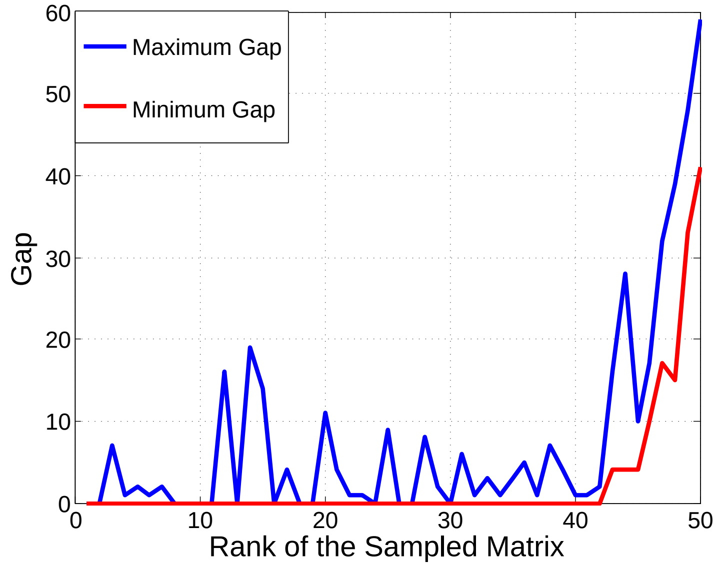

Figure 3: The gaps between the rank of the obtained matrix via (17) and that of the original sampled matrix for $p = 0.46$.

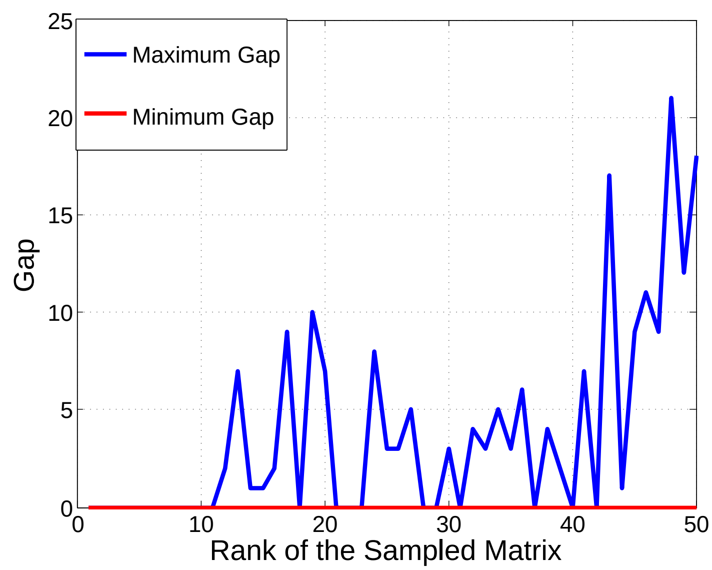

Figure 4: The gaps between the rank of the obtained matrix via (17) and that of the original sampled matrix for $p = 0.50$.

- As we can observe in Figs. 3–6, the defined gap is always a nonnegative number, which is consistent with previous observation that for $p = 0.54$ and $p = 0.58$ we can certify that with high probability (≥ 0.6) the rank of any completion is an upper bound on the rank of the sampled matrix or r^*.

- For $p = 0.54$ and $p = 0.58$ that we have theoretical results (as mentioned in the first observation) the gap obtained by (17) is very close to zero. This phenomenon (that we do not have a rigorous justification for) shows that as soon as we can certify our proposed theoretical

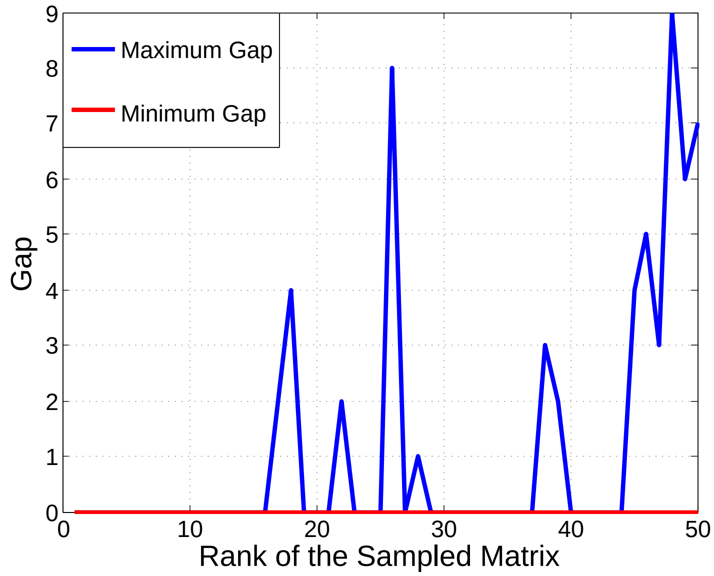

Figure 5: The gaps between the rank of the obtained matrix via (17) and that of the original sampled matrix for $p = 0.54$.

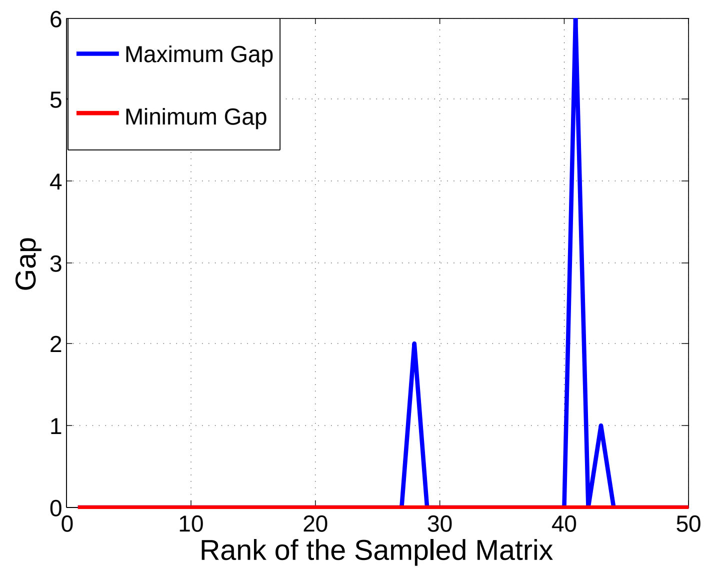

Figure 6: The gaps between the rank of the obtained matrix via (17) and that of the original sampled matrix for $p = 0.58$.

results (i.e., as soon as the rank of a completion provides an upper bound on the rank of the sampled matrix or r^*) by increasing the sampling probability, the upper bound found through (17) becomes very tight; in some cases this bound is exactly equal to r^* (red curves) and in some cases this bound is almost equal to r^* (blue curves). However, these gaps are not small (specially blue curves) for $p = 0.46$ and $p = 0.50$ and note that according to Fig. 1, for these values of p we cannot guarantee the bounds on the value of rank hold with high probability.

3.2 CP-Rank Tensor

Let $\mathbb{P}_i$ denote the Lebesgue measures on $\mathbb{R}^{n_i \times r^*}$, $i = 1, \ldots, d$. In this subsection, we assume that the sampled tensor $\mathcal{U} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ is chosen generically from the manifold of tensors of rank $r^* = \text{rank}_{\text{CP}}(\mathcal{U})$ (where r^* is unknown), or in other words, the entries of $\mathcal{U}$ are drawn independently with respect to Lebesgue measure on the corresponding manifold. Hence, the probability measures of all statements in this subsection are $\mathbb{P}_1 \times \mathbb{P}_2 \times \ldots \times \mathbb{P}_d$.

Condition $\mathcal{A}_r$: Each row of the d-th matricization of the sampled tensor, i.e., $\mathbf{U}_{(d)}$ includes at least r observed entries.

We construct a binary valued tensor called **constraint tensor** $\breve{\Omega}_r$ based on Ω and a given number r. Consider any subtensor $\mathcal{Y} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_{d-1} \times 1}$ of the tensor $\mathcal{U}$. The sampled tensor $\mathcal{U}$ includes n_d subtensors that belong to $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_{d-1} \times 1}$ and let $\mathcal{Y}_i$ for $1 \leq i \leq n_d$ denote these n_d subtensors. Define a binary valued tensor $\breve{\mathcal{Y}}_i \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_{d-1} \times k_i}$, where $k_i = N_\Omega(\mathcal{Y}_i) - r$ and its entries are described as the following. We can look at $\breve{\mathcal{Y}}_i$ as k_i tensors each belongs to $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_{d-1} \times 1}$. For each of the mentioned k_i tensors in $\breve{\mathcal{Y}}_i$ we set the entries corresponding to r of the observed entries equal to 1. For each of the other k_i observed entries, we pick one of the k_i tensors of $\breve{\mathcal{Y}}_i$ 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. In the case that $k_i = 0$ we simply ignore $\breve{\mathcal{Y}}_i$, i.e., $\breve{\mathcal{Y}}_i = \emptyset$

By putting together all n_d tensors in dimension d, we construct a binary valued tensor $\breve{\Omega}_r \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_{d-1} \times K}$, where $K = \sum_{i=1}^{n_d} k_i = N_\Omega(\mathcal{U}) - r n_d$ and call it the **constraint tensor**. Observe that each subtensor of $\breve{\Omega}_r$ which belongs to $\mathbb{R}^{n_1 \times n_2 \times \cdots \times n_{d-1} \times 1}$ includes exactly $r + 1$ nonzero entries. In (Ashraphijuo and Wang, 2017b), an example is given on the construction of $\breve{\Omega}_r$.

Condition $\mathcal{B}_r$: $\breve{\Omega}_r$ consists a subtensor $\breve{\Omega}'_r \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_{d-1} \times K}$ such that $K = r(\sum_{i=1}^{d-1} n_i) - r^2 - r(d-2)$ and for any $K' \in \{1, 2, \ldots, K\}$ and any subtensor $\breve{\Omega}''_r \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_{d-1} \times K'}$ of $\breve{\Omega}'_r$ we have

$$r \left(\left(\sum_{i=1}^{d-1} f_i(\breve{\Omega}''_r) \right) - \min \left\{ \max \left\{ f_1(\breve{\Omega}''_r), \ldots, f_{d-1}(\breve{\Omega}''_r) \right\}, r \right\} - (d-2) \right) \geq K', \tag{18}$$

where $f_i(\breve{\Omega}''_r)$ denotes the number of nonzero rows of the i-th matricization of $\breve{\Omega}''_r$.

The following lemma is a re-statement of Theorem 1 in (Ashraphijuo and Wang, 2017b).

Lemma 13 *With probability one, there are only finitely many rank-r^* completions of the sampled tensor if and only if Conditions $\mathcal{A}_{r^*}$ and $\mathcal{B}_{r^*}$ hold.*

Definition 14 *Let $\mathcal{S}_\Omega$ denote the set of all natural numbers r such that both Conditions $\mathcal{A}_r$ and $\mathcal{B}_r$ hold.*

Lemma 15 *There exists a number r_Ω such that $\mathcal{S}_\Omega = \{1, 2, \ldots, r_\Omega\}$.*

Proof The proof is similar to the proof of Lemma 3 since the dimension of the manifold of CP rank-r tensors is $r(\sum_{i=1}^{d} n_i) - r^2 - r(d-1)$, which is an increasing function in r. ■

The following theorem gives an upper bound on the unknown rank r^*.

Theorem 16 *With probability one, exactly one of the following statements holds*

(i) $r^* \in \mathcal{S}_\Omega = \{1, 2, \ldots, r_\Omega\}$;

(ii) For any arbitrary completion of the sampled tensor $\mathcal{U}$ of rank r, we have $r \notin \mathcal{S}_\Omega$.

Proof Similar to the proof of Theorem 4, it suffices to show that the assumption $r^* \notin \mathcal{S}_\Omega$ results that there exists a completion of $\mathcal{U}$ of CP rank r, where $r \in \mathcal{S}_\Omega$, with probability zero. Define $\mathcal{V} = (\mathcal{V}_1, \dots, \mathcal{V}_r)$ as the basis of the rank-r CP decomposition of $\mathcal{U}$ as in (3), where $\mathcal{V}_l = \mathbf{a}_1^l \otimes \mathbf{a}_2^l \otimes \dots \otimes \mathbf{a}_{d-1}^l \in \mathbb{R}^{n_1 \times \dots n_{d-1}}$ is a rank-1 tensor and $\mathbf{a}_i^l$ is defined in (3) for $1 \leq l \leq r$ and $1 \leq i \leq d$. Define $\mathcal{Y} = (\mathbf{a}_d^1, \dots, \mathbf{a}_d^r)$ and $\mathcal{V} \otimes_d \mathcal{Y} = \sum_{l=1}^r \mathcal{V}_l \otimes \mathbf{a}_d^l$. Observe that $\mathcal{U} = \sum_{l=1}^r \mathcal{V}_l \otimes \mathbf{a}_d^l = \mathcal{V} \otimes_d \mathcal{Y}$.

Observe that each row of $\mathbf{U}_{(d)}$ includes at least $r_\Omega + 1$ observed entries since Condition $\mathcal{A}_{r_\Omega}$ holds. Moreover, the existence of a completion of the sampled tensor $\mathcal{U}$ of rank $r \in \mathcal{S}_\Omega$ results in the existence of a basis $\mathcal{V} = (\mathcal{V}_1, \dots, \mathcal{V}_r)$ such that there exists $\mathcal{Y} = (\mathbf{a}_d^1, \dots, \mathbf{a}_d^r)$ and $\mathcal{U}_\Omega = (\mathcal{V} \otimes_d \mathcal{Y})_\Omega$. As a result, given $\mathcal{V}$, each observed entry of $\mathcal{U}$ results in a degree-1 polynomial in terms of the entries of $\mathcal{Y}$ as

$$\mathcal{U}(\vec{x}) = \sum_{l=1}^{r} \mathcal{V}_l(x_1, \dots, x_{d-1}) \mathbf{a}_d^l(x_d). \tag{19}$$

Note that $r_\Omega \geq r$ and each row of $\mathbf{U}_{(d)}$ includes at least $r_\Omega + 1 \geq r + 1$ observed entries. Consider $r+1$ of the observed entries of the first row of $\mathbf{U}_{(d)}$ and we denote them by $\mathcal{U}(\vec{x}_1)$, $\dots$, $\mathcal{U}(\vec{x}_{r+1})$, where the last component of the vector $\vec{x}_i$ is equal to one, $1 \leq i \leq r+1$. Similar to the proof of Theorem 4, genericity of $\mathcal{U}$ results in

$$\mathcal{U}(\vec{x}_{r+1}) = \sum_{l=1}^{r} t_l \mathcal{U}(\vec{x}_i), \tag{20}$$

where t_l's are constant scalars, $l = 1, \dots, r$. On the other hand, according to Lemma 13 there exist at most finitely many completions of the sampled tensor of rank r. Therefore, there exist a completion of $\mathbf{U}$ of rank r with probability zero. Moreover, an equation of the form of (20) holds with probability zero as $r^* \geq r + 1$ and $\mathcal{U}$ is chosen generically from the manifold of tensors of rank-r^*. Therefore, there exists a completion of rank r with probability zero. ■

Corollary 17 *Consider an arbitrary number $r' \in \mathcal{S}_\Omega$. Similar to Theorem 16, it follows that with probability one, exactly one of the followings holds*

(i) $r^* \in \{1, 2, \dots, r'\}$;

(ii) For any arbitrary completion of the sampled tensor $\mathcal{U}$ of rank r, we have $r \notin \{1, 2, \dots, r'\}$.

Corollary 18 *Assuming that there exists a CP rank-r completion of the sampled tensor $\mathcal{U}$ such that $r \in \mathcal{S}_\Omega$, we conclude that with probability one $r^* \leq r$.*

Corollary 19 *Let $\mathcal{U}^*$ denote an optimal solution to the following NP-hard optimization problem*

$$\begin{aligned} &\text{minimize}_{\mathcal{U}' \in \mathbb{R}^{n_1 \times \cdots \times n_d}} && \text{rank}_{\text{CP}}(\mathcal{U}') \\ &\text{subject to} && \mathcal{U}'_\Omega = \mathcal{U}_\Omega. \end{aligned} \tag{21}$$

Assume that $\text{rank}_{\text{CP}}(\mathcal{U}^*) \in \mathcal{S}_\Omega$. *Then, Corollary 18 results that* $r^* = \text{rank}_{\text{CP}}(\mathcal{U}^*)$ *with probability one.*

The following lemma is Lemma 15 in (Ashraphijuo and Wang, 2017b), which is the probabilistic version of Lemma 13 in terms of the sampling probability.

Lemma 20 *Assume that* $n_1 = n_2 = \cdots = n_d = n$, $d > 2$, $n > \max\{200, r(d-2)\}$ *and* $r \leq \frac{n}{6}$. *Moreover, assume that the sampling probability satisfies*

$$p > \frac{1}{n^{d-2}} \max\left\{27 \log\left(\frac{n}{\epsilon}\right) + 9 \log\left(\frac{2r(d-2)}{\epsilon}\right) + 18, 6r\right\} + \frac{1}{\sqrt[4]{n^{d-2}}}. \tag{22}$$

Then, with probability at least $(1-\epsilon)\left(1 - \exp(-\frac{\sqrt{n^{d-2}}}{2})\right)^{n^2}$, *we have* $r \in \mathcal{S}_\Omega$.

The following corollary is the probabilistic version of Corollaries 18 and 19.

Corollary 21 *Assuming that there exists a CP rank-*r *completion of the sampled tensor* $\mathcal{U}$ *such that the conditions given in Lemma 20 hold, with the sampling probability satisfying* (22)*, we conclude that with probability at least* $(1-\epsilon)\left(1 - \exp(-\frac{\sqrt{n^{d-2}}}{2})\right)^{n^2}$ *we have* $r^* \leq r$. *Therefore, given that* (22) *holds for* $r = \text{rank}(\mathbf{U}^*)$ *and* $\mathbf{U}^*$ *denotes an optimal solution to the optimization problem* (21)*, with probability at least* $(1-\epsilon)\left(1 - \exp(-\frac{\sqrt{n^{d-2}}}{2})\right)^{n^2}$ *we have* $r^* = \text{rank}(\mathbf{U}^*)$.

3.2.1 NUMERICAL RESULTS

We generate a random tensor $\mathcal{U} \in \mathbb{R}^{8\times8\times8\times8\times8\times8}$ of CP-rank 2 by adding two random rank-1 tensors. The color scale represents the lower bound on the probability that we can guarantee the rank of a given completion is an upper bound on the true value of rank. Then, we solve the following convex optimization problem for different values of the sampling probability.

$$\begin{array}{ll} \text{minimize}_{\mathcal{U}' \in \mathbb{R}^{n_1\times\cdots\times n_d}} & \|\widetilde{\mathbf{U}}'_{(3)}\|_* \\ \text{subject to} & \mathcal{U}'_\Omega = \mathcal{U}_\Omega. \end{array} \tag{23}$$

Note that rank of any of the unfoldings of a tensor is a lower bound on the CP-rank of that tensor. Hence, we minimize the nuclear norm of the unfolding with the possible maximum rank among all unfoldings as $\widetilde{\mathbf{U}}_{(3)} \in \mathbb{R}^{512\times512}$. Then, we use the Matlab toolbox found online "Tensorlab" to calculate the CP-rank of the obtained tensor via solving convex program (23) (there are other methods to calculate CP decomposition, e.g., (Pimentel-Alarcón, 2016)). In Figure 7, gap represents the CP-rank of the solution of (23) minus the CP-rank of the original sampled tensor.

4. Vector-Rank Cases

4.1 Multi-View Matrix

Let $\mathbb{P}_1$ and $\mathbb{P}_2$ denote the Lebesgue measures on $\mathbb{R}^{n\times r_1^*}$ and $\mathbb{R}^{r_1^*\times n_1}$, respectively. Moreover, let $\mathbb{P}_3$ and $\mathbb{P}_4$ denote the Lebesgue measures on $\mathbb{R}^{n\times(r^*-r_1^*)}$ and $\mathbb{R}^{r_2^*\times n_2}$, respectively. In this paper, we assume that $\mathbf{U}$ is chosen generically from the manifold corresponding to rank vector (r_1^*, r_2^*, r^*), i.e., the entries of $\mathbf{U}$ are drawn independently with respect to Lebesgue measure on the corresponding manifold. Hence, the probability measures of all statements in this subsection are $\mathbb{P}_1\times\mathbb{P}_2\times\mathbb{P}_3\times\mathbb{P}_4$.

The following Conditions will be used frequently in this subsection.

Condition A_{r_1,r_2}: Each column of $\mathbf{U}_1$ and $\mathbf{U}_2$ include at least r_1 and r_2 sampled entries, respectively.

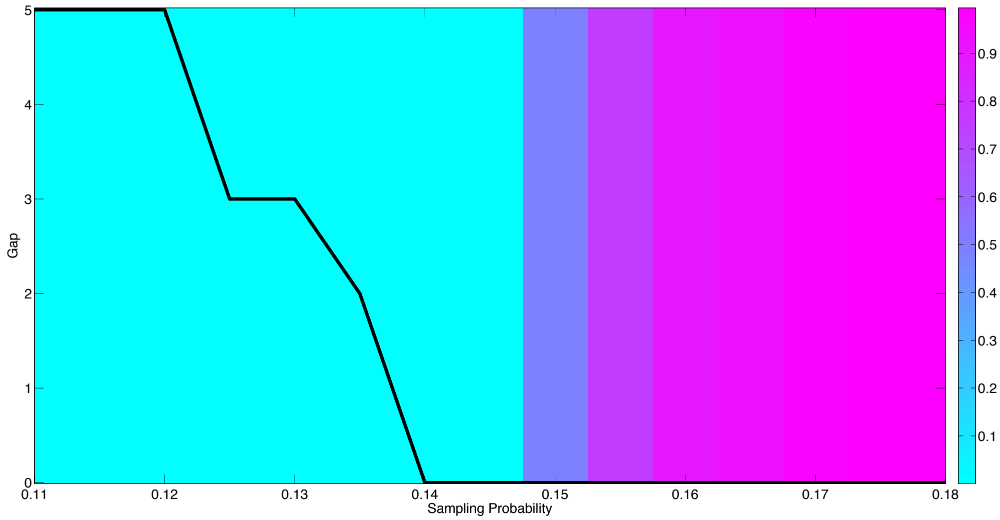

Figure 7: The rank gap as a function of sampling probability for $\mathcal{U} \in \mathbb{R}^{8\times8\times8\times8\times8\times8}$ of CP-rank 2.

We construct a binary valued matrix called **constraint matrix** for multi-view matrix $\mathbf{U}$ as $\breve{\mathbf{\Omega}}_{r_1,r_2} = [\breve{\mathbf{\Omega}}_{r_1}|\breve{\mathbf{\Omega}}_{r_2}]$, where $\breve{\mathbf{\Omega}}_{r_1}$ and $\breve{\mathbf{\Omega}}_{r_2}$ represent the constraint matrix for single-view matrices $\mathbf{U}_1$ and $\mathbf{U}_2$ (defined in Section 3.1), respectively.

Condition $B_{r_1,r_2,r}$: $\breve{\mathbf{\Omega}}_{r_1,r_2}$ consists a submatrix $\breve{\mathbf{\Omega}}'_{r_1,r_2} \in \mathbb{R}^{n\times K}$ such that $K = nr - r^2 - r_1^2 - r_2^2 + r(r_1 + r_2)$ and for any $K' \in \{1, 2, \dots, K\}$ and any submatrix $\breve{\mathbf{\Omega}}''_{r_1,r_2} \in \mathbb{R}^{n\times K'}$ of $\breve{\mathbf{\Omega}}'_{r_1,r_2}$ we have

$$(r - r_2)\left(f(\breve{\mathbf{\Omega}}''_{r_1}) - r_1\right)^+ + (r - r_1)\left(f(\breve{\mathbf{\Omega}}''_{r_2}) - r_2\right)^+ \\ +(r_1 + r_2 - r)\left(f(\breve{\mathbf{\Omega}}''_{r_1,r_2}) - (r_1 + r_2 - r)\right)^+ \geq K', \quad (24)$$

where $f(\mathbf{X})$ denotes the number of nonzero rows of $\mathbf{X}$ for any matrix $\mathbf{X}$ and $\breve{\mathbf{\Omega}}''_{r_1,r_2} = [\breve{\mathbf{\Omega}}''_{r_1}|\breve{\mathbf{\Omega}}''_{r_2}]$, and also $\breve{\mathbf{\Omega}}''_{r_1}$ and $\breve{\mathbf{\Omega}}''_{r_2}$ denote the columns of $\breve{\mathbf{\Omega}}''_{r_1,r_2}$ corresponding to $\breve{\mathbf{\Omega}}_{r_1}$ and $\breve{\mathbf{\Omega}}_{r_2}$, respectively.

The following lemma is a re-statement of Theorem 2 in (Ashraphijuo et al., 2017c).

Lemma 22 *With probability one, there are only finitely many completions of the sampled multi-view data if and only if Conditions $A_{r_1^*,r_2^*}$ and $B_{r_1^*,r_2^*,r^*}$ hold.*

Definition 23 *Denote the rank vector $\underline{r} = (r_1, r_2, r)$. Define the generalized inequality $\underline{r}' \preceq \underline{r}$ as the component-wise set of inequalities, e.g., $r'_1 \leq r_1$, $r'_2 \leq r_2$ and $r' \leq r$.*

Definition 24 *Let $\mathcal{S}_{\mathbf{\Omega}}$ denote the set of all $\underline{r}$ such that both Conditions A_{r_1,r_2} and $B_{r_1,r_2,r}$ hold.*

Lemma 25 *Assume $\underline{r} \in \mathcal{S}_{\mathbf{\Omega}}$. Then, for any $\underline{r}' \preceq \underline{r}$, we have $\underline{r}' \in \mathcal{S}_{\mathbf{\Omega}}$.*

Proof We consider the rank factorization of $\mathbf{U}$ as in (Ashraphijuo et al., 2017c) and similar to the single-view scenario in Lemma 3 each observed entry results in a polynomial in terms of the entries of the components of the decomposition. Note that the dimension of the manifold corresponding to rank vector $\underline{r}$ is equal to $rn + r_1n_1 + r_2n_2 - r^2 - r_1^2 - r_2^2 + r(r_1 + r_2)$ (Ashraphijuo et al., 2017c), and also observe that the fact that $\max\{r_1, r_2\} \leq r \leq r_1+r_2 \leq \min\{2n, n_1+n_2\}$ implies that reducing any of the values r_1, r_2, and r reduces the value of $rn + r_1n_1 + r_2n_2 - r^2 - r_1^2 - r_2^2 + r(r_1 + r_2)$. Hence, the dimension of the manifold corresponding to rank vector $\underline{r}$ is larger than that for rank vector $\underline{r}'$, given $\underline{r}' \preceq \underline{r}$, and thus similar to the proof of Lemma 3, finite completability of data with $\underline{r}$ results finite completability of data with $\underline{r}'$ with probability one. Then, using Lemma 22, the proof is complete. ■

The following theorem provides a relationship between the unknown rank vector $\underline{r}^*$ and $\mathcal{S}_{\mathbf{\Omega}}$.

Theorem 26 *With probability one, exactly one of the following statements holds*

(i) $\underline{r}^* \in \mathcal{S}_{\mathbf{\Omega}}$;

(ii) For any arbitrary completion of the sampled matrix $\mathbf{U}$ *of rank vector* $\underline{r}$, *we have* $\underline{r} \notin \mathcal{S}_{\mathbf{\Omega}}$.

Proof Similar to the proof of Theorem 4, suppose that there does not exist a completion of $\mathbf{U}$ of rank vector $\underline{r}$ such that $\underline{r} \in \mathcal{S}_{\mathbf{\Omega}}$. Therefore, it is easily verified that statement (ii) holds and statement (i) does not hold. On the other hand, assume that there exists a completion of $\mathbf{U}$ of rank vector $\underline{r}$, where $\underline{r} \in \mathcal{S}_{\mathbf{\Omega}}$. Hence, statement (ii) does not hold and to complete the proof it suffices to show that with probability one, statement (i) holds. Similar to Theorem 4, we show that assuming $\underline{r}^* \notin \mathcal{S}_{\mathbf{\Omega}}$, there exists a completion of $\mathbf{U}$ of rank vector $\underline{r}$, where $\underline{r} \in \mathcal{S}_{\mathbf{\Omega}}$, with probability zero.

Since $\underline{r}^* \notin \mathcal{S}_{\mathbf{\Omega}}$, according to Lemma 25, for any $\underline{r} \in \mathcal{S}_{\mathbf{\Omega}}$ at least one the following inequalities holds; $r_1 < r_1^*$, $r_2 < r_2^*$ and $r < r^*$. Note that assuming that there exists a completion of $\mathbf{U}_1$ of rank r_1 with probability zero results that there exists a completion of $\mathbf{U}$ of rank vector $\underline{r}$ with probability zero and similar statement holds for r_2 and r. Hence, in any possible scenario ($r_1 < r_1^*$ or $r_2 < r_2^*$ or $r < r^*$) the similar proof as in Theorem 4 (for single-view matrix) results that there exists a completion of $\mathbf{U}$ of rank vector $\underline{r}$, where $\underline{r} \in \mathcal{S}_{\mathbf{\Omega}}$, with probability zero. ■

Corollary 27 *Consider a subset* $\mathcal{S}'_{\mathbf{\Omega}}$ *of* $\mathcal{S}_{\mathbf{\Omega}}$ *such that for any two members of* $\mathcal{S}_{\mathbf{\Omega}}$ *that* $\underline{r}' \preceq \underline{r}''$ *and* $\underline{r}'' \in \mathcal{S}'_{\mathbf{\Omega}}$ *we have* $\underline{r}' \in \mathcal{S}'_{\mathbf{\Omega}}$. *Then, with probability one, exactly one of the followings holds*

(i) $\underline{r}^* \in \mathcal{S}'_{\mathbf{\Omega}}$;

(ii) For any arbitrary completion of $\mathbf{U}$ *of rank vector* $\underline{r}$, *we have* $\underline{r} \notin \mathcal{S}'_{\mathbf{\Omega}}$.

Proof Note that the property in the statement of Lemma 25 holds for $\mathcal{S}'_{\mathbf{\Omega}}$ as well as $\mathcal{S}_{\mathbf{\Omega}}$. Moreover, as $\mathcal{S}'_{\mathbf{\Omega}} \subseteq \mathcal{S}_{\mathbf{\Omega}}$, for any $\underline{r} \in \mathcal{S}'_{\mathbf{\Omega}}$ there exists at most finitely many completions of $\mathbf{U}$ of rank vector $\underline{r}$, and therefore the rest of the proof is the same as the proof of Theorem 26. ■

Corollary 28 *Assuming that there exists a completion of* $\mathbf{U}$ *with rank vector* $\underline{r}$ *such that* $\underline{r} \in \mathcal{S}_{\mathbf{\Omega}}$, *then with probability one* $\underline{r}^* \preceq \underline{r}$.

The following lemma which is a re-statement of Theorem 3 in (Ashraphijuo et al., 2017c) gives the number of samples per column that is needed to ensure that Conditions A_{r_1,r_2} and $B_{r_1,r_2,r}$ hold with high probability.

Lemma 29 *Suppose that the following inequalities hold*

$$\begin{aligned} \frac{n}{6} &\geq \max\{r_1, r_2, (r_1+r_2-r)\}, & (25)\\ n_1 &\geq (r-r_2)(n-r_1), & (26)\\ n_2 &\geq (r-r_1)(n-r_2), & (27)\\ n_1+n_2 &\geq (r-r_2)(n-r_1)+(r-r_1)(n-r_2) \\ &+ (r_1+r_2-r)(n-(r_1+r_2-r)). & (28) \end{aligned}$$

Moreover assume that each **column** *of* $\mathbf{U}$ *is observed in at least* l *entries, uniformly at random and independently across entries, where*

$$l > \max\left\{9\ \log\left(\frac{n}{\epsilon}\right) + 3\ \log\left(\frac{3\max\{r-r_2, r-r_1, r_1+r_2-r\}}{\epsilon}\right) + 6, 2r_1, 2r_2\right\}. \quad (29)$$

Then, with probability at least $1-\epsilon$, $\underline{r} \in \mathcal{S}_{\mathbf{\Omega}}$.

The following proposition is the probabilistic version of Theorem 26 in terms of the sampling probability instead of verifying Conditions A_{r_1,r_2} and $B_{r_1,r_2,r}$.

Proposition 30 *Suppose that* (25)-(28) *hold for* $\underline{r}$ *and that each entry of the sampled matrix is observed uniformly at random and independently across entries with probability* p*, where*

$$p > \frac{1}{n}\max\left\{9\ \log\left(\frac{n}{\epsilon}\right) + 3\ \log\left(\frac{3\max\{r-r_2, r-r_1, r_1+r_2-r\}}{\epsilon}\right) + 6, 2r_1, 2r_2\right\} + \frac{1}{\sqrt[4]{n}}.$$

Then, with probability at least $(1-\epsilon)\left(1-\exp(-\frac{\sqrt{n}}{2})\right)^{n_1+n_2}$*, we have* $\underline{r} \in \mathcal{S}_{\mathbf{\Omega}}$.

Proof The proposition is easy to verify using Lemma 29 and Lemma 9 (similar to the proof for Proposition 11). ■

Corollary 31 *Assuming that there exists a completion of* $\mathbf{U}$ *of rank vector* $\underline{r}$ *such that* (25)-(28) *hold and the sampling probability satisfies* (30)*, then with probability at least* $(1-\epsilon)\left(1-\exp(-\frac{\sqrt{n}}{2})\right)^{n_1+n_2}$ *we have* $\underline{r}^* \preceq \underline{r}$.

4.2 Tucker-Rank Tensor

Let $\mathbb{P}_i$ denote the Lebesgue measures on $\mathbb{R}^{n_i \times m_i^*}$, $i = j+1, \dots, d$, and $\mathbb{P}_0$ denotes the Lebesgue measures on $\mathbb{R}^{m_{j+1}^* \times m_{j+2}^* \times \dots \times m_d^*}$. In this subsection, we assume that the sampled tensor $\mathcal{U} \in \mathbb{R}^{n_1 \times \dots \times n_d}$ is chosen generically from the manifold of tensors of rank $\underline{r}^* = \text{rank}_{\text{Tucker}}(\mathcal{U}) = (m_{j+1}^*, \dots, m_d^*)$ (where $\underline{r}^*$ is unknown), or in other words, the entries of $\mathcal{U}$ are drawn independently with respect to Lebesgue measure on the corresponding manifold. Hence, the probability measures of all statements in this subsection are $\mathbb{P}_0 \times \mathbb{P}_{j+1} \times \mathbb{P}_{j+2} \times \dots \times \mathbb{P}_d$.

Without loss of generality assume that $m_{j+1}^* \geq \dots \geq m_d^*$ throughout this subsection. Also, given $\underline{r} = (m_{j+1}, \dots, m_d)$, define the following function

$$g_{\underline{r}}(x) = \sum_{i=j+1}^{d} \min\left\{r_i, \left(x - \sum_{i'=j+1}^{i-1} r_{i'}\right)^+\right\} r_i. \quad (30)$$

Definition 32 *For any $i \in \{j+1,\dots,d\}$ and $\mathcal{S}_i \subseteq \{1,\dots,n_i\}$, define $\mathcal{U}^{(\mathcal{S}_i)}$ as a set containing the entries of $|\mathcal{S}_i|$ rows (corresponding to the elements of $\mathcal{S}_i$) of $\mathbf{U}_{(i)}$. Moreover, define $\mathcal{U}^{(\mathcal{S}_{j+1},\dots,\mathcal{S}_d)} = \mathcal{U}^{(\mathcal{S}_{j+1})} \cup \dots \cup \mathcal{U}^{(\mathcal{S}_d)}$.*

Condition $\mathcal{A}_{\underline{r}}^{\text{Tucker}}$: There exist $\sum_{i=j+1}^{d} (n_i m_i)$ observed entries such that for any $\mathcal{S}_i \subseteq \{1,\dots, n_i\}$ for $i \in \{j+1,\dots,d\}$, $\mathcal{U}^{(\mathcal{S}_{j+1},\dots,\mathcal{S}_d)}$ includes at most $\sum_{i=j+1}^{d} |\mathcal{S}_i| m_i$ of the mentioned $\sum_{i=j+1}^{d} n_i m_i$ observed entries.

Let $\mathcal{P}$ be a set of $\sum_{i=j+1}^{d} (n_i m_i)$ observed entries such that they satisfy Condition $\mathcal{A}_{\underline{r}}^{\text{Tucker}}$. Now, we construct a $(j+1)^{\text{th}}$-order binary constraint tensor $\breve{\Omega}_{\underline{r}}$ in some sense similar to that in Section 3.2. 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 $\mathcal{U}$, let $N_\Omega(\mathcal{Y}^{\mathcal{P}})$ denote the number of sampled entries in $\mathcal{Y}$ that belong to $\mathcal{P}$.

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 $\mathcal{Y}_{(t_{j+1},\dots,t_d)}$ where $(t_{j+1},\dots,t_d)$ represents the coordinate of the subtensor. Define a binary valued tensor $\breve{\mathcal{Y}}_{(t_{j+1},\cdots,t_d)} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times \overbrace{1 \times \dots \times 1}^{d-j} \times k}$, where $k = N_\Omega(\mathcal{Y}_{(t_{j+1},\dots,t_d)}) - N_\Omega(\mathcal{Y}^{\mathcal{P}}_{(t_{j+1},\dots,t_d)})$ 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 $N_\Omega(\mathcal{Y}^{\mathcal{P}}_{(t_{j+1},\dots,t_d)})$ observed entries that belong to $\mathcal{P}$ 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 $\breve{\Omega}_{\underline{r}} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times K_j}$, where $K_j = N_\Omega(\mathcal{U}) - \sum_{i=j+1}^{d} (n_i m_i)$ and call it the **constraint tensor** (Ashraphijuo et al., 2016a). In (Ashraphijuo et al., 2016a), an example is given on the construction of $\breve{\Omega}_{\underline{r}}$.

Condition $\mathcal{B}_{\underline{r}}^{\text{Tucker}}$: The constraint tensor $\breve{\Omega}_{\underline{r}}$ consists a subtensor $\breve{\Omega}'_{\underline{r}} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_j \times K}$ such that $K = \left(\Pi_{i=1}^{j} n_i\right)\left(\Pi_{i=j+1}^{d} m_i\right) - \sum_{i=j+1}^{d} m_i^2$ and for any $K' \in \{1,2,\dots,K\}$ and any subtensor $\breve{\Omega}''_{\underline{r}} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_{d-1} \times K'}$ of $\breve{\Omega}'_{\underline{r}}$ we have

$$\left(\Pi_{i=j+1}^{d} m_i\right)\left(f_{j+1}(\breve{\Omega}''_{\underline{r}})\right) - g_{\underline{r}}\left(f_{j+1}(\breve{\Omega}''_{\underline{r}})\right) \geq K', \tag{31}$$

where $f_{j+1}(\breve{\Omega}''_{\underline{r}})$ denotes the number of nonzero columns of the $(j+1)$-th matricization of $\breve{\Omega}''_{\underline{r}}$.

The following lemma is a re-statement of Theorem 3 in (Ashraphijuo et al., 2016a).

Lemma 33 *With probability one, there are only finitely many completions of rank $\underline{r}^*$ of the sampled tensor if and only if Conditions $\mathcal{A}_{\underline{r}^*}^{\text{Tucker}}$ and $\mathcal{B}_{\underline{r}^*}^{\text{Tucker}}$ hold.*

Definition 34 *Let $\mathcal{S}_\Omega$ denote the set of all rank vectors $\underline{r}$ such that both Conditions $\mathcal{A}_{\underline{r}}^{\text{Tucker}}$ and $\mathcal{B}_{\underline{r}}^{\text{Tucker}}$ hold.*

Lemma 35 *Assume $\underline{r} \in \mathcal{S}_\Omega$. Then, for any rank vector $\underline{r}' \preceq \underline{r}$, we have $\underline{r}' \in \mathcal{S}_\Omega$.*

Proof Note that the dimension of the manifold corresponding to $\underline{r}$ is $\left(\Pi_{i=1}^{j} n_i\right)\left(\Pi_{i=j+1}^{d} m_i\right) + \sum_{i=j+1}^{d} n_i m_i - \sum_{i=j+1}^{d} m_i^2$, and thus by reducing the value of m_{i_0} by one (for $i_0 \in \{j+1, \dots, d\}$), the value of the mentioned dimension reduces by at least $\left(\Pi_{i=1}^{j} n_i\right) + n_i - 2m_i + 1$, which is greater than zero since $m_i \leq n_i$. The rest of the proof is similar to the proof of Lemma 3. ■

Definition 36 *Define $\mathcal{S}_\Omega(\underline{r})$ as a subset of $\mathcal{S}_\Omega$, which includes all $\underline{r}' \in \mathcal{S}_\Omega$ that $\underline{r}' \preceq \underline{r}$.*

The following theorem gives a relationship between $\underline{r}^*$ and $\mathcal{S}_\Omega$.

Theorem 37 *With probability one, exactly one of the following statements holds*
(i) $\underline{r}^* \in \mathcal{S}_\Omega$;
(ii) For any arbitrary completion of the sampled tensor $\mathcal{U}$ of rank $\underline{r}$, we have $\underline{r} \notin \mathcal{S}_\Omega(\underline{r}^*)$.

Proof Similar to the proof of Theorem 4, to complete the proof it suffices to show that the assumption $\underline{r}^* \notin \mathcal{S}_\Omega$ results that there exists a completion of $\mathcal{U}$ of rank $\underline{r}$, where $\underline{r} \in \mathcal{S}_\Omega(\underline{r}^*)$, with probability zero. Note that $\underline{r} \in \mathcal{S}_\Omega(\underline{r}^*) \subseteq \mathcal{S}_\Omega$ results that Conditions $\mathcal{A}_{\underline{r}}^{\text{Tucker}}$ and $\mathcal{B}_{\underline{r}}^{\text{Tucker}}$ hold. Moreover, note that $\underline{r} \preceq \underline{r}^*$ and since $\underline{r}^* \notin \mathcal{S}_\Omega$ we conclude that there exists $i_0 \in \{j+1, \dots, d\}$ such that $m_{i_0} < m_{i_0}^*$. As a result, $\sum_{i=j+1}^{d} n_i m_i < \sum_{i=j+1}^{d} n_i m_i^*$.

Condition $\mathcal{B}_{\underline{r}}^{\text{Tucker}}$ ensures there exists at least one more observed entry (otherwise the constraint tensor does not exist) besides the $\sum_{i=j+1}^{d} n_i m_i$ mentioned observed entries. Given the basis $\mathcal{C} \in \mathbb{R}^{n_1 \times \dots \times n_j \times m_{j+1} \times \dots \times m_d}$ as in (7), there exist $\sum_{i=j+1}^{d} n_i m_i$ variables in the corresponding Tucker decomposition. However, we have $\sum_{i=j+1}^{d} n_i m_i + 1$ polynomials in terms these $\sum_{i=j+1}^{d} n_i m_i$ variables and therefore the last polynomials can be written as algebraic combination of the other $\sum_{i=j+1}^{d} n_i m_i$ polynomials. This leads to a linear equation in terms of the $\sum_{i=j+1}^{d} n_i m_i + 1$ corresponding observed entries. On the other hand, the $\sum_{i=j+1}^{d} n_i m_i$ observed entries satisfy the property stated as Condition $\mathcal{A}_{\underline{r}}^{\text{Tucker}}$ and it is easily verified that there exist $\sum_{i=j+1}^{d} n_i m_i^*$ entries (observed and non-observed) satisfying Condition $\mathcal{A}_{\underline{r}^*}^{\text{Tucker}}$ such that the union of the mentioned $\sum_{i=j+1}^{d} n_i m_i$ entries with any arbitrary other observed entry be a subset of those $\sum_{i=j+1}^{d} n_i m_i^*$ entries. However, $\mathcal{U}$ is generically chosen from the manifold corresponding to $\underline{r}^*$ and therefore a particular linear equation in terms of the mentioned $\sum_{i=j+1}^{d} n_i m_i^*$ entries holds with probability zero. The rest of the proof is similar to the proof of Theorem 4. ■

Corollary 38 *Assuming that there exists a completion of $\mathcal{U}$ with rank vector $\underline{r}$ such that $\underline{r} \in \mathcal{S}_\Omega$, we conclude that with probability one $\underline{r}^* \preceq \underline{r}$.*

The following lemma is Corollary 2 in (Ashraphijuo et al., 2016a), which ensures that Conditions $\mathcal{A}_{\underline{r}}^{\text{Tucker}}$ and $\mathcal{B}_{\underline{r}}^{\text{Tucker}}$ hold with high probability.

Lemma 39 *Assume that $\sum_{i=j+1}^{d} m_i^2 \leq \Pi_{i=j+1}^{d} m_i$, $\Pi_{i=j+1}^{d} n_i \geq N_j \Pi_{i=j+1}^{d} m_i - \sum_{i=j+1}^{d} m_i^2$, $\Pi_{i=j+1}^{d} m_i \leq N_j$, where $N_j = \Pi_{i=1}^{j} n_i$. Furthermore, assume that we observe each entry of $\mathcal{U}$ with* *probability* p*, where*

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

Then, with probability at least $(1-\epsilon)\left(1-\exp(-\frac{\sqrt{\Pi_{i=1}^{j} n_i}}{2})\right)^{\Pi_{i=j+1}^{d} n_i}$, $\underline{r} \in \mathcal{S}_\Omega$.

The following corollary is the probabilistic version of Theorem 37.

Corollary 40 *Assuming that there exists a completion of the sampled tensor* $\mathcal{U}$ *of Tucker rank* $\underline{r}$ *such that the assumptions in Lemma 39 hold and the sampling probability satisfies* (32)*, then with probability at least* $(1-\epsilon)\left(1-\exp(-\frac{\sqrt{\Pi_{i=1}^{j} n_i}}{2})\right)^{\Pi_{i=j+1}^{d} n_i}$ *we have* $\underline{r}^* \preceq \underline{r}$*.*

4.2.1 NUMERICAL RESULTS

We generate a random tensor $\mathcal{U} \in \mathbb{R}^{8\times 8\times 8\times 8\times 8\times 8}$ of Tucker-rank $(1,3,3,2,2)$. The color scale represents the lower bound on the probability that we can guarantee the rank of a given completion is a component-wise upper bound on the true rank. Then, we solve the following convex optimization problem for different values of the sampling probability.

$$\begin{aligned} &\text{minimize}_{\mathcal{U}' \in \mathbb{R}^{n_1\times\cdots\times n_d}} && \|\sum_{i=1}^{d} \mathbf{U}'_{(i)}\|_* \\ &\text{subject to} && \mathcal{U}'_\Omega = \mathcal{U}_\Omega. \end{aligned} \tag{32}$$

Then, we calculate rank of each matricization of the tensor obtained via solving (32) to find its Tucker-rank. In this scenario, for each component of the Tucker-rank, we find the percentage of error via $\frac{m_i - m_i^*}{n - m_i^*} \times 100\%$, where $n = 8$, m_i and m_i^* are the i-th rank component of the obtained tensor and original tensor, respectively. Hence, 100% error simply means that the corresponding matricization is full rank. In Figure 8, gap represents the average of the defined error over all components of Tucker-rank, i.e., over all matricizations.

4.3 TT-Rank Tensor

Let $\mathbb{P}_i$ denote the Lebesgue measures on $\mathbb{R}^{u_{i-1}^*\times n_i \times u_i^*}$, $i = 1,\dots,d$, where $u_0^* = u_d^* = 1$. In this subsection, we assume that the sampled tensor $\mathcal{U} \in \mathbb{R}^{n_1\times\cdots\times n_d}$ is chosen generically from the manifold of tensors of rank $\underline{r}^* = \text{rank}_{\text{TT}}(\mathcal{U}) = (u_1^*,\dots,u_{d-1}^*)$ (where $\underline{r}^*$ is unknown), or in other words, the entries of $\mathcal{U}$ are drawn independently with respect to Lebesgue measure on the corresponding manifold. Hence, the probability measures of all statements in this subsection are $\mathbb{P}_1 \times \ldots \times \mathbb{P}_d$.

Condition $\mathcal{A}_{\underline{r}}^{\text{TT}}$: Each row of the d-th matricization of the sampled tensor, i.e., $\mathbf{U}_{(d)}$ includes at least u_{d-1} observed entries.

We construct the d-way binary valued constraint tensor $\breve{\Omega}_{u_{d-1}}$ similar to that in Section 3.2 as the following. Consider any subtensor $\mathcal{Y} \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_{d-1}\times 1}$ of the tensor $\mathcal{U}$. The sampled tensor

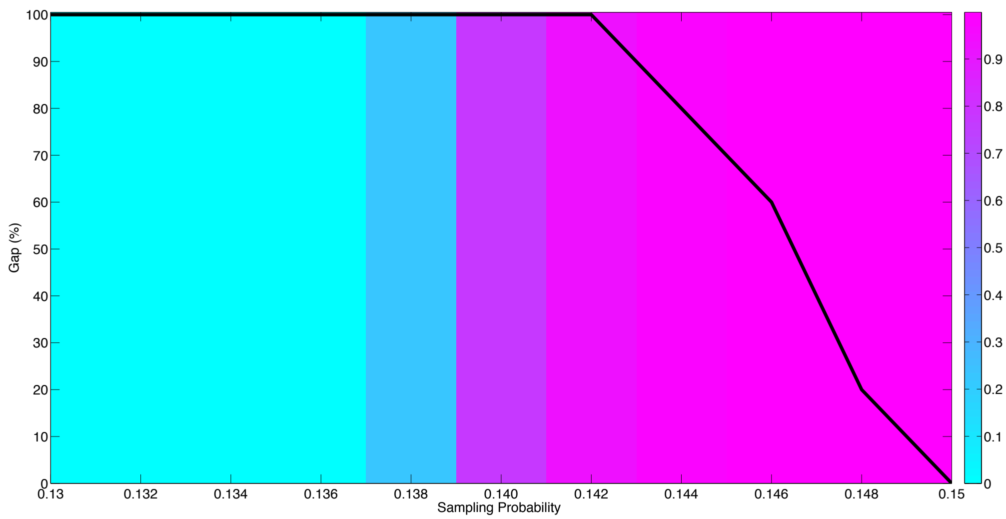

Figure 8: The rank gap as a function of sampling probability for $\mathcal{U} \in \mathbb{R}^{8\times 8\times 8\times 8\times 8\times 8}$ of Tucker-rank $(1, 3, 3, 2, 2)$.

$\mathcal{U}$ includes n_d subtensors that belong to $\mathbb{R}^{n_1\times n_2\times\cdots\times n_{d-1}\times 1}$ and let $\mathcal{Y}_i$ for $1 \leq i \leq n_d$ denote these n_d subtensors. Define a binary valued tensor $\breve{\mathcal{Y}}_i \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_{d-1}\times k_i}$, where $k_i = N_\Omega(\mathcal{Y}_i) - u_{d-1}$ and its entries are described as the following. We can look at $\breve{\mathcal{Y}}_i$ as k_i tensors each belongs to $\mathbb{R}^{n_1\times n_2\times\cdots\times n_{d-1}\times 1}$. For each of the mentioned k_i tensors in $\breve{\mathcal{Y}}_i$ we set the entries corresponding to u_{d-1} of the observed entries equal to 1. For each of the other k_i observed entries, we pick one of the k_i tensors of $\breve{\mathcal{Y}}_i$ 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. In the case that $k_i = 0$ we simply ignore $\breve{\mathcal{Y}}_i$, i.e., $\breve{\mathcal{Y}}_i = \emptyset$

By putting together all n_d tensors in dimension d, we construct a binary valued tensor $\breve{\Omega}_{u_{d-1}} \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_{d-1}\times K}$, where $K = \sum_{i=1}^{n_d} k_i = N_\Omega(\mathcal{U}) - u_{d-1}n_d$ and call it the **constraint tensor**. Observe that each subtensor of $\breve{\Omega}_{u_{d-1}}$ which belongs to $\mathbb{R}^{n_1\times n_2\times\cdots\times n_{d-1}\times 1}$ includes exactly $u_{d-1}+1$ nonzero entries. In (Ashraphijuo and Wang, 2017a), an example is given on the construction of $\breve{\Omega}_{u_{d-1}}$.

Condition $\mathcal{B}_{\underline{r}}^{\mathrm{TT}}$: $\breve{\Omega}_{u_{d-1}}$ consists a subtensor $\breve{\Omega}'_{u_{d-1}} \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_{d-1}\times K}$ such that $K = \sum_{i=1}^{d-1} u_{i-1}n_iu_i - \sum_{i=1}^{d-1} u_i^2$ and for any $K' \in \{1, 2, \ldots, K\}$ and any subtensor $\breve{\Omega}''_{u_{d-1}} \in \mathbb{R}^{n_1\times n_2\times\cdots\times n_{d-1}\times K'}$ of $\breve{\Omega}'_{u_{d-1}}$ we have

$$\sum_{i=1}^{d-1} \left(u_{i-1} f_i(\breve{\Omega}''_{u_{d-1}}) u_i - u_i^2 \right)^+ \geq K', \tag{33}$$

where $f_i(\breve{\Omega}''_{u_{d-1}})$ denotes the number of nonzero rows of the i-th matricization of $\breve{\Omega}''_{u_{d-1}}$.

The following lemma is a re-statement of Theorem 1 in (Ashraphijuo and Wang, 2017a).

Lemma 41 *With probability one, there are only finitely many completions of rank $\underline{r}^*$ of the sampled tensor if and only if Conditions $\mathcal{A}_{\underline{r}^*}^{TT}$ and $\mathcal{B}_{\underline{r}^*}^{TT}$ hold.*

Definition 42 *Let $\mathcal{S}_\Omega$ denote the set of all rank vectors $\underline{r}$ such that both Conditions $\mathcal{A}_{\underline{r}}^{TT}$ and $\mathcal{B}_{\underline{r}}^{TT}$ hold.*

The following lemma will be used in Lemma 44.

Lemma 43 $u_i \leq \min\{u_{i-1}n_i, u_{i+1}n_{i+1}\}$ *for* $1 \leq i \leq d-1$.

Proof We first show that $u_i \leq u_{i-1}n_i$, which is easily verified for $i = 1$ as $\widetilde{\mathbf{U}}_1$ includes n_1 rows and $u_0 = 1$, and therefore assume that $i > 1$. Define the $(d-1)$-way tensor $\mathcal{U}^{l_i} \in \mathbb{R}^{n_1 \times \ldots \times n_{i-1} \times n_{i+1} \times \ldots \times n_d}$ such that $\mathcal{U}^{l_i}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d) = \mathcal{U}(x_1, \ldots, x_{i-1}, l_i, x_{i+1}, \ldots, x_d)$ for $1 \leq i \leq d$ and $1 \leq l_i \leq n_i$. Also, recall that $\widetilde{\mathbf{U}}^{l_i}_{(i-1)}$ denotes the $(i-1)$-th unfolding of $\mathcal{U}^{l_i}$. Observe that $\widetilde{\mathbf{U}}^{l_i}_{(i-1)}$ is a subset of columns of matrix $\widetilde{\mathbf{U}}_{(i-1)}$ (those columns that correspond to the entries of $\mathcal{U}$ with the i-th component of the location equal to l_i). Therefore, rank$(\widetilde{\mathbf{U}}^{l_i}_{(i-1)}) \leq$ rank$(\widetilde{\mathbf{U}}_{(i-1)}) = u_{i-1}$.

On the other hand, observe that $\widetilde{\mathbf{U}}^{l_i}_{(i-1)}$ is a subset of rows of $\widetilde{\mathbf{U}}_{(i)}$ (those rows that correspond to the entries of $\mathcal{U}$ with the i-th component of the location equal to l_i). Hence, the union of rows of $\widetilde{\mathbf{U}}^{l_i}_{(i-1)}$'s for $1 \leq l_i \leq n_i$ constitute all rows of $\widetilde{\mathbf{U}}_{(i)}$. Therefore, $u_i =$ rank$(\widetilde{\mathbf{U}}_{(i)}) \leq \sum_{l_i=1}^{n_i}$ rank$(\widetilde{\mathbf{U}}^{l_i}_{(i-1)}) \leq n_i u_{i-1}$. Similarly, we can show that $u_i \leq u_{i+1}n_{i+1}$ to complete the proof. ■

Lemma 44 *Assume* $\underline{r} \in \mathcal{S}_\Omega$*. Then, for any* $\underline{r}' \preceq \underline{r}$*, we have* $\underline{r}' \in \mathcal{S}_\Omega$*.*

Proof Note that the dimension of the manifold corresponding to $\underline{r}$ is $\sum_{i=1}^{d} u_{i-1}n_iu_i - \sum_{i=1}^{d-1} u_i^2$. If we reduce the value of u_i by one, the value of the mentioned dimension reduces by $u_{i-1}n_i + u_{i+1}n_{i+1} - 2u_i + 1$. According to Lemma 43, $u_{i-1}n_i + u_{i+1}n_{i+1} - 2u_i + 1$ is greater than zero, and therefore $\underline{r}' \preceq \underline{r}$ results that the dimension of the manifold corresponding to $\underline{r}$ is greater than that corresponding to $\underline{r}'$. The rest of the proof is similar to the proof of Lemma 3. ■

Definition 45 *Define* $\hat{\mathcal{S}}_\Omega$ *as the set of all rank vectors* $\underline{r} \in \mathcal{S}_\Omega$ *such that there exists a rank vector* $\underline{r}' \in \mathcal{S}_\Omega$ *with* $\underline{r} \preceq \underline{r}'$ *and* $u_{d-1} < u'_{d-1}$ *(instead of* $u_{d-1} \leq u'_{d-1}$*). Note that* $\hat{\mathcal{S}}_\Omega$ *also satisfies the property in Lemma 44.*

Theorem 46 *With probability one, exactly one of the following statements holds:*
(i) $\underline{r}^* \in \hat{\mathcal{S}}_\Omega$;
(ii) For any arbitrary completion of the sampled tensor $\mathcal{U}$ *of rank* $\underline{r}$*, we have* $\underline{r} \notin \hat{\mathcal{S}}_\Omega$*.*

Proof Similar to the proof of Theorem 4, to complete the proof it suffices to show that the assumption $\underline{r}^* \notin \hat{\mathcal{S}}_\Omega$ results that there exists a completion of $\mathcal{U}$ of rank $\underline{r}$, where $\underline{r} \in \hat{\mathcal{S}}_\Omega$, with probability zero. Define the multiplication $\mathcal{U}^{(1)} \ldots \mathcal{U}^{(d-1)}$ in (9) as the basis of the rank $\underline{r}$ TT decomposition of $\mathcal{U}$. Then, by considering the $(d-1)$-th unfolding of $\mathcal{U}^{(1)} \ldots \mathcal{U}^{(d-1)}$ in TT decomposition we obtain a matrix factorization of the $(d-1)$-th unfolding of $\mathcal{U}$. The rest of the proof is similar to the proof of Theorem 4. ■

Similar to Theorem 46, we can show the following.

Corollary 47 *Consider a subset $\hat{\mathcal{S}}'_\Omega$ of $\hat{\mathcal{S}}_\Omega$ such that for any two members of $\hat{\mathcal{S}}_\Omega$ that $\underline{r}'' \preceq \underline{r}'$ and $\underline{r}' \in \hat{\mathcal{S}}'_\Omega$ we have $\underline{r}'' \in \hat{\mathcal{S}}'_\Omega$. Then, with probability one, exactly one of the followings holds*

(i) $\underline{r}^* \in \hat{\mathcal{S}}'_\Omega$;

(ii) For any arbitrary completion of $\mathcal{U}$ of rank vector $\underline{r}$, we have $\underline{r} \notin \hat{\mathcal{S}}'_\Omega$.

Corollary 48 *Assuming that there exists a completion of $\mathcal{U}$ with rank vector $\underline{r}$ such that $\underline{r} \in \hat{\mathcal{S}}_\Omega$, we conclude that with probability one $\underline{r}^* \preceq \underline{r}$.*

The following lemma is Lemma 14 in (Ashraphijuo and Wang, 2017a), which ensures that Conditions $\mathcal{A}_{\underline{r}}^{\text{TT}}$ and $\mathcal{B}_{\underline{r}}^{\text{TT}}$ hold with high probability.

Lemma 49 *Define $m = \sum_{k=1}^{d-2} u_{k-1}u_k$, $M = n\sum_{k=1}^{d-2} u_{k-1}u_k - \sum_{k=1}^{d-2} u_k^2$ and $u' = \max\left\{\frac{u_1}{u_0}, \dots, \frac{u_{d-2}}{u_{d-3}}\right\}$. Assume that $n_1 = n_2 = \cdots = n_d = n$, $n > \max\{m, 200\}$ and $u' \leq \min\{\frac{n}{6}, u_{d-2}\}$ hold. Moreover, assume that the sampling probability satisfies*

$$p > \frac{1}{n^{d-2}} \max\left\{27 \ \log\left(\frac{n}{\epsilon}\right) + 9 \ \log\left(\frac{2M}{\epsilon}\right) + 18, 6u_{d-2}\right\} + \frac{1}{\sqrt[4]{n^{d-2}}}. \tag{34}$$

Then, with probability at least $(1-\epsilon)\left(1 - \exp(-\frac{\sqrt{n^{d-2}}}{2})\right)^{n^2}$, we have $\underline{r} \in \mathcal{S}_\Omega$.

The following corollary is the probabilistic version of Corollary 48.

Corollary 50 *Assuming that there exists a completion of the sampled tensor $\mathcal{U}$ of TT rank $\underline{r}$ such that the assumptions in Lemma 49 hold and the sampling probability satisfies* (34)*, then with probability at least $(1-\epsilon)\left(1 - \exp(-\frac{\sqrt{n^{d-2}}}{2})\right)^{n^2}$ we have $\underline{r}^* \preceq \underline{r}$.*

4.3.1 NUMERICAL RESULTS

We generate a random tensor $\mathcal{U} \in \mathbb{R}^{8\times8\times8\times8\times8\times8}$ of TT-rank $(1,2,4,1,1)$. The color scale represents the lower bound on the probability that we can guarantee the rank of a given completion is a component-wise upper bound on the true rank. Then, we solve the following convex optimization problem for different values of the sampling probability.

$$\begin{aligned} &\text{minimize}_{\mathcal{U}' \in \mathbb{R}^{n_1\times\cdots\times n_d}} && \|\sum_{i=1}^{d-1} \widetilde{\mathbf{U}}'_{(i)}\|_* \\ &\text{subject to} && \mathcal{U}'_\Omega = \mathcal{U}_\Omega. \end{aligned} \tag{35}$$

Then, we calculate rank of each unfolding of the tensor obtained via solving (35) to find its TT-rank. In this scenario, for each component of the TT-rank, we find the percentage of error via $\frac{u_i - u_i^*}{\min\{n^i, n^{d-i}\} - u_i^*} \times 100\%$, where $n = 8$, $d = 6$, u_i and u_i^* are the i-th rank component of the obtained tensor and original tensor, respectively. Hence, 100% error simply means that the corresponding unfolding is full rank. In Figure 9, gap represents the average of the defined error over all components of TT-rank, i.e., over all unfoldings.

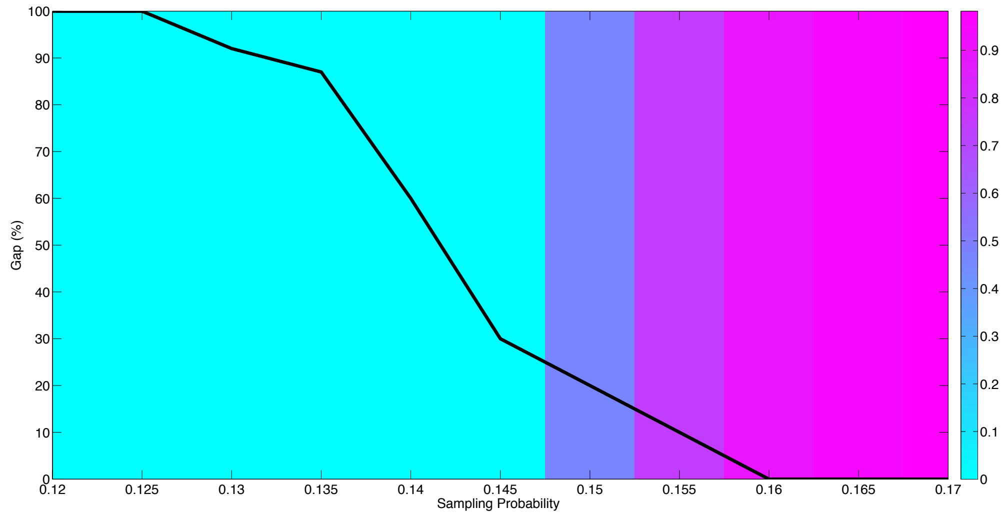

Figure 9: The rank gap as a function of sampling probability for $\mathcal{U} \in \mathbb{R}^{8\times8\times8\times8\times8\times8}$ of TT-rank $(1, 2, 4, 1, 1)$.

5. Conclusions

We make use of the recently developed algebraic geometry analyses that study the fundamental conditions on the sampling patterns for finite completability under a number of low-rank matrix and tensor models to treat the problem of rank approximation for a partially sampled data. Particularly, the goal is to approximate the unknown scalar or vector rank based on the sampling pattern and the rank of a given completion. A number of data models have been treated, including single-view matrix, multi-view matrix, CP tensor, tensor-train tensor and Tucker tensor. First we have provided an upper bound on the unknown scalar rank (for single-view matrix and CP tensor) and an component-wise upper bound on the vector rank (for multi-view matrix, Tucker tensor and TT tensor) with probability one assuming that the sampling pattern satisfies the proposed combinatorial conditions. Moreover, we have also provided probabilistic versions of such bounds that hold with high probability assuming that the sampling probability is above a threshold. In addition, for single-view matrix and CP tensor, these upper bounds can be exactly equal to the unknown scalar rank given the lowest-rank completion. To illustrate how tight our proposed upper bounds are, we have provided some numerical results for the single-view matrix case in which we applied the nuclear norm minimization to find a low-rank completion of the sampled data and observe that the proposed upper bound is almost equal to the true unknown rank.

Acknowledgments

This work was supported in part by the U.S. National Science Foundation (NSF) under grant CIF1064575, and in part by the U.S. Office of Naval Research (ONR) under grant N000141410667.

References

Morteza Ashraphijuo and Xiaodong Wang. Characterization of deterministic and probabilistic sampling patterns for finite completability of low tensor-train rank tensor. *arXiv preprint:1703.07698*, 2017a.

Morteza Ashraphijuo and Xiaodong Wang. Fundamental conditions for low-CP-rank tensor completion. *Journal of Machine Learning Research*, 18 (63):1–29, 2017b.

Morteza Ashraphijuo and Xiaodong Wang. Scaled nuclear norm minimization for low-rank tensor completion. *arXiv preprint:1707.07976*, 2017c.

Morteza Ashraphijuo, Ramtin Madani, and Javad Lavaei. Inverse function theorem for polynomial equations using semidefinite programming. In *IEEE 54th Annual Conference on Decision and Control (CDC)*, pages 6589–6596, 2015.

Morteza Ashraphijuo, Vaneet Aggarwal, and Xiaodong Wang. Deterministic and probabilistic conditions for finite completability of low rank tensor. *arXiv preprint:1612.01597*, 2016a.

Morteza Ashraphijuo, Ramtin Madani, and Javad Lavaei. Characterization of rank-constrained feasibility problems via a finite number of convex programs. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 6544–6550, 2016b.

Morteza Ashraphijuo, Ramtin Madani, and Javad Lavaei. Characterization of rank-constrained feasibility problems via a finite number of convex programs. In *IEEE 55th Conference on Decision and Control (CDC)*, pages 6544–6550, 2016c.

Morteza Ashraphijuo, Vaneet Aggarwal, and Xiaodong Wang. A characterization of sampling patterns for low-Tucker-rank tensor completion problem. In *IEEE International Symposium on Information Theory (ISIT)*, pages 531–535, 2017a.

Morteza Ashraphijuo, Xiaodong Wang, and Vaneet Aggarwal. A characterization of sampling patterns for low-rank multi-view data completion problem. In *IEEE International Symposium on Information Theory (ISIT)*, pages 1147–1151, 2017b.

Morteza Ashraphijuo, Xiaodong Wang, and Vaneet Aggarwal. Deterministic and probabilistic conditions for finite completability of low-rank multi-view data. *arXiv preprint:1701.00737*, 2017c.

Morteza Ashraphijuo, Xiaodong Wang, and Vaneet Aggarwal. An approximation of the CP-rank of a partially sampled tensor. In *55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2017d.

Jian Feng Cai, Emmanuel J Candès, and Zuowei Shen. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, 20(4):1956–1982, 2010.

Emmanuel Candes and Benjamin Recht. Exact matrix completion via convex optimization. *Communications of the ACM*, 55(6):111–119, 2012.

Emmanuel J Candès and Benjamin Recht. Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, 9(6):717–772, 2009.

Emmanuel J Candès and Terence Tao. The power of convex relaxation: Near-optimal matrix completion. *IEEE Transactions on Information Theory*, 56(5):2053–2080, 2010.

Emmanuel J Candès, Yonina C Eldar, Thomas Strohmer, and Vladislav Voroninski. Phase retrieval via matrix completion. *SIAM Journal on Imaging Sciences*, 6(1):199–225, 2013.

Lars Eldén. *Matrix Methods in Data Mining and Pattern Recognition*, volume 4. Society for Industrial and Applied Mathematics, 2007.

Silvia Gandy, Benjamin Recht, and Isao Yamada. Tensor completion and low-n-rank tensor recovery via convex optimization. *Inverse Problems*, 27(2):1–19, 2011.

Rong Ge, Jason D Lee, and Tengyu Ma. Matrix completion has no spurious local minimum. *arXiv preprint:1605.07272*, 2016.

Nicholas JA Harvey, David R Karger, and Kazuo Murota. Deterministic network coding by matrix completion. In *the annual ACM-SIAM symposium on discrete algorithms*, pages 489–498, 2005.

Prateek Jain, Praneeth Netrapalli, and Sujay Sanghavi. Low-rank matrix completion using alternating minimization. In *Annual Symposium on the Theory of Computing*, pages 665–674, 2013.

Hui Ji, Chaoqiang Liu, Zuowei Shen, and Yuhong Xu. Robust video denoising using low rank matrix completion. In *CVPR*, pages 1791–1798, 2010.

Lek-Heng Lim and Pierre Comon. Multiarray signal processing: Tensor decomposition meets compressed sensing. *Comptes Rendus Mecanique*, 338(6):311–320, 2010.

Ji Liu, Przemyslaw Musialski, Peter Wonka, and Jieping Ye. Tensor completion for estimating missing values in visual data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):208–220, 2013.

X. Y. Liu, S. Aeron, V. Aggarwal, X. Wang, and M. Y. Wu. Adaptive sampling of RF fingerprints for fine-grained indoor localization. *IEEE Transactions on Mobile Computing*, 15(10):2411–2423, 2016.

Oyetunji E Ogundijo, Abdulkadir Elmas, and Xiaodong Wang. Supplemental material for reverse engineering gene regulatory networks from measurement with missing values.

Oyetunji E Ogundijo, Abdulkadir Elmas, and Xiaodong Wang. Reverse engineering gene regulatory networks from measurement with missing values. *EURASIP Journal on Bioinformatics and Systems Biology*, 2017(1):2, 2017.

D Pimentel-Alarcón. A simpler approach to low-rank tensor canonical polyadic decomposition. In *54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 474–481. IEEE, 2016.

D Pimentel-Alarcón, Nigel Boston, and Robert D Nowak. Deterministic conditions for subspace identifiability from incomplete sampling. In *IEEE International Symposium on Information Theory (ISIT)*, pages 2191–2195, 2015.

D Pimentel-Alarcón, L Balzano, R Marcia, R Nowak, and R Willett. Group-sparse subspace clustering with missing data. In *IEEE Workshop on Statistical Signal Processing (SSP)*, pages 1–5, 2016a.

D Pimentel-Alarcón, EDU Robert D Nowak, and WISC EDU. The information-theoretic requirements of subspace clustering with missing data. In *International Conference on Machine Learning*, 2016b.

Daniel Pimentel-Alarcón and Robert Nowak. A converse to low-rank matrix completion. In *IEEE International Symposium on Information Theory (ISIT)*, 2016.

Daniel Pimentel-Alarcón, Laura Balzano, and Robert Nowak. Necessary and sufficient conditions for sketched subspace clustering. In *Allerton Conference on Communication, Control, and Computing*, 2016c.

Daniel Pimentel-Alarcón, Nigel Boston, and Robert Nowak. A characterization of deterministic sampling patterns for low-rank matrix completion. *IEEE Journal of Selected Topics in Signal Processing*, 10(4):623–636, 2016d.

Benjamin Recht and Christopher Ré. Parallel stochastic gradient algorithms for large-scale matrix completion. *Mathematical Programming Computation*, 5(2):201–226, 2013.

Gil Shabat. Matrix completion using nuclear norm, spectral norm or weighted nuclear norm minimization. 2015.

Nicholas D Sidiropoulos and Anastasios Kyrillidis. Multi-way compressed sensing for sparse low-rank tensors. *IEEE Signal Processing Letters*, 19(11):757–760, 2012.