Pattern Recognition 103 (2020) 107307

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/patcog

Fundamental sampling patterns for low-rank multi-view data completion☆

Morteza Ashraphijuo[a], Xiaodong Wang[a,*], Vaneet Aggarwal[b]

[a] Department of Electrical Engineering, Columbia University, NY, United States
[b] School of Industrial Engineering, Purdue University, West Lafayette, IN, United States

ARTICLE INFO

Article history:
Received 8 April 2019
Revised 16 December 2019
Accepted 24 February 2020
Available online 29 February 2020

Keywords:
Multi-view learning
Low-rank matrix completion
Sampling pattern
Sampling rate
Non-convex optimization
Rank decomposition

ABSTRACT

We consider the multi-view data completion problem, i.e., to complete a matrix $\mathbf{U} = [\mathbf{U}_1|\mathbf{U}_2]$ where the ranks of $\mathbf{U}, \mathbf{U}_1$, and $\mathbf{U}_2$ are given. In particular, we investigate the fundamental conditions on the sampling pattern, i.e., locations of the sampled entries for finite completability of such a multi-view data given the corresponding rank constraints. We provide a geometric analysis on the manifold structure for multi-view data to incorporate more than one rank constraint. We derive a probabilistic condition in terms of the number of samples per column that guarantees finite completability with high probability. Finally, we derive the guarantees for unique completability. Numerical results demonstrate reduced sampling complexity when the multi-view structure is taken into account as compared to when only low-rank structure of individual views is taken into account. Then, we propose an apporach using Newton's method to almost achieve these information-theoretic bounds for mulit-view data retrieval by taking advantage of the rank decomposition and the analysis in this work.

1. Introduction

High-dimensional data analysis has received significant recent attention due to the ubiquitous big data, including images and videos, product ranking datasets, gene expression database, etc. Many real-world high-dimensional datasets exhibit low-rank structures, i.e., the data can be represented in a much lower dimensional form [1]. Efficiently exploiting such low-rank structure for analyzing large high-dimensional datasets is one of the most active research area in machine learning and data mining. The data structure that we study in this work is multi-view data with two views that are individually low rank and the concatenated matrix formed by the two views is also low rank. With this multi-view structure, this paper aims to complete the data in the presence of missing entries.

With one view, the problem is a standard low-rank matrix completion problem which has been widely studied [2,3]. For the model of multi-view, each view can be considered separately as a matrix completion problem. However, the low rank assumption of the concatenation of the two views provide additional structure, and thus the requirement of the data may decrease. Exploring the sampling requirements for this multi-view problem is the focus of this paper. Multi-view learning problem has applications in signal processing [4], multi-label image classification [5–7], data clustering [8], image retrieval [9], image synthesis [10,11], data classification [12], rank estimation [13], multi-lingual text categorization [14], etc.

We assume that the ranks of each view and the concatenated data are provided, and the data at the sampled entries are known. Using this information, a completion is any matrix that agrees with the sampled entries and rank constraints. The multi-view data is called finitely completable if and only if there exist only finitely many completions. For a single view (matrix completion), multiple optimization-based methods have been proposed including alternating minimization [15], convex relaxation of rank [16–18], etc. Moreover, there are many optimization-based analyses in the literature for multi-view learning [19–29]. Moreover, the multi-view clustering problem using matrix decompsotion-based approaches is studied in [30,31] and other subspace analysis-based methods in [32,33].

The optimization-based matrix completion algorithms typically require incoherence conditions, which constrains the values of the entries (sampled and non-sampled) to obtain a completion with high probability. Moreover, the fundamental completability conditions that are independent of the specific completion algorithms have also been investigated. Specifically, deterministic conditions

☆ The deterministic analysis part of this work is partially presented in ISIT 2017, Aachen, Germany.

* Corresponding author.
E-mail addresses: ashraphijuo@ee.columbia.edu (M. Ashraphijuo), wangx@ee.columbia.edu (X. Wang), vaneet@purdue.edu (V. Aggarwal).

https://doi.org/10.1016/j.patcog.2020.107307
0031-3203/

on the locations of the sampled entries (sampling pattern) are obtained through algebraic geometry analyses on Grassmannian manifold that lead to finite/unique solutions to the matrix completion problem [34–36]. The analysis in [34] provides the combinatorial conditions on the location of the sampled entries for finite/unique completability of the sampled matrix of the given rank. Such an algorithm-independent condition can lead to a much lower sampling rate than the one that is required by the optimization-based completion algorithms. However, the analysis on Grassmannian manifold in [34] is not capable of incorporating more than one rank constraint. Even though the analysis for matrices have been extended to tensors in [37,38], that extension also do not incorporate rank constraints for the different views. As the first step, we study the finite completability problem for multi-view data by proposing a geometrical analysis on the manifold structure for such data. Moreover, other interesting related problems have been studied using algebraic geometry analysis, including high-rank matrix completion [39] and subspace clustering with missing data [40–43].

This paper aims to provide the lower bounds on the number of sampled entries per column such that the proposed conditions on $\mathbf{\Omega}$ for finite/unique completability are satisfied with high probability. This work is inspired by [34], where the analysis on Grassmannian manifold is proposed to solve similar problems for a single-view matrix. Specifically, in [34] a novel approach is proposed to consider the rank factorization of a matrix and to treat each observed entry as a polynomial in terms of the entries of the components of the rank factorization. Then, under the genericity assumption, the algebraic independence among the mentioned polynomials is studied. For the multi-view data completion problem, we first follow the general approach that is similar to that in [34] to treat the corresponding problem. Although these works have similar natures, they are fundamentally different. The fact that we consider the rank decomposition of the data that corresponds to multiple rank constraints instead of one rank constraint, results in different polynomial structure for each sampled entry. We mention some of the main differences of our approach as compared to that in [34]. The different geometry of the manifold leads to a change in the geometry of the manifold structure. This further leads to a difference in the equivalence class for the basis and the canonical basis. This further leads to a different structure on the algebraic polynomials which makes the analysis not directly extendable. Moreover, the idea of using all three rank constraints simultaneously in the algebraic geometry approach has not been considered so far. Hence, the manifold structure for the multi-view data is fundamentally different from the Grassmannian manifold and we need to develop almost every step anew.

The rest of the paper is organized as follows. In Section 2, the notations and problem statement are provided. In Section 4, the guarantees for finite completability are proposed where the condition is in terms of the number of samples per column. In Section 5, the guarantees for unique completability are provided. Numerical results are provided in Section 6 to compare the number of samples per column for finite and unique completions based on our proposed analysis versus the existing method. We see that exploiting the additional structure provided by multi-view model leads to a significant decrease in the number of samples required for data completion. Sections Appendix A–Appendix B provides a proof for finite completability of example in Section 2 and describes a larger example, respectively. Finally, Section 8 concludes the paper.

2. Problem statement and a motivating example

Let $\mathbf{U}$ be the sampled data that we want to study. Denote $\mathbf{\Omega}$ as the sampling pattern matrix that is of the same size as $\mathbf{U}$ and $\mathbf{\Omega}(x_1, x_2) = 1$ if $\mathbf{U}(x_1, x_2)$ is observed and $\mathbf{\Omega}(x_1, x_2) = 0$ otherwise. For each subset of columns $\mathbf{U}'$ of $\mathbf{U}$, define $N_{\mathbf{\Omega}}(\mathbf{U}')$ as the number of observed entries in $\mathbf{U}'$ according to the sampling pattern $\mathbf{\Omega}$. For any real number x, define $x^+ = \max\{0, x\}$. Also, $\mathbf{I}_n$ denotes an $n \times n$ identity matrix and $\mathbf{0}_{n \times m}$ denotes an $n \times m$ all-zero matrix.

The matrix $\mathbf{U} \in \mathbb{R}^{n\times(m_1+m_2)}$ is randomly sampled. Denote a partition of $\mathbf{U}$ as $\mathbf{U} = [\mathbf{U}_1|\mathbf{U}_2]$ where $\mathbf{U}_1 \in \mathbb{R}^{n\times m_1}$ and $\mathbf{U}_2 \in \mathbb{R}^{n\times m_2}$ represent the first and second views of data, respectively. Given the rank constraints $\text{rank}(\mathbf{U}_1) = r_1$, $\text{rank}(\mathbf{U}_2) = r_2$ and $\text{rank}(\mathbf{U}) = r$, our goal is to characterize the geometrical conditions on the locations of the sampled entries to ensure there exist infinite, finite, or unique completions of the sampled data $\mathbf{U}$ such that given rank constraints hold true.

In [34], a necessary and sufficient condition on the sampling pattern is given for the finite completability of a matrix $\mathbf{U}$ given $\text{rank}(\mathbf{U}) = r$, based on an algebraic geometry analysis on the Grassmannian manifold. However, this analysis cannot be used to incorporate the three rank constraints for the multi-view problem since we have multiple rank constraints here. This is because the rank decomposition corresponding to three rank constraints has a different nature and as we will see this would change the whole analysis.

3. Geometry of the basis

In this section, we will define an equivalence relation among all bases of the sampled matrix $\mathbf{U}$, where a basis is a set of r vectors ($r = \text{rank}(\mathbf{U})$) that spans the column space of $\mathbf{U}$. This equivalence relation leads to the manifold structure for multi-view data to incorporate all three rank constraints.

First, we provide an example such that: (i) given r_1, $\mathbf{U}_1$ is infinitely completable; (ii) given r_2, $\mathbf{U}_2$ is infinitely completable; (iii) given r, $\mathbf{U}$ is infinitely completable; and (iv) given r_1, r_2 and r, $\mathbf{U}$ is finitely completable. This is equivalent to showing that applying the single-view analysis to $\mathbf{U}$, $\mathbf{U}_1$, and $\mathbf{U}_2$ separately does not guaranty the finite completability of any of $\mathbf{U}$, $\mathbf{U}_1$, and $\mathbf{U}_2$, but incorporating all three rank constraints in the multi-view analysis guaranties the finite completability. In other words, if $\mathcal{S}_1$ denotes the set of completions of $\mathbf{U}$ given $\text{rank}(\mathbf{U}_1) = r_1$, $\mathcal{S}_2$ denotes the set of completions of $\mathbf{U}$ given $\text{rank}(\mathbf{U}_2) = r_2$ and $\mathcal{S}$ denotes the set of completions of $\mathbf{U}$ given $\text{rank}(\mathbf{U}) = r$, then in the following example $|\mathcal{S}_1| = \infty$, $|\mathcal{S}_2| = \infty$, $|\mathcal{S}| = \infty$ and $|\mathcal{S}_1 \cap \mathcal{S}_2 \cap \mathcal{S}| < \infty$.

Before explaining the example, we provide a simple lemma, which will be used in the example.

Lemma 1. *Consider a sampled rank-r matrix* $\mathbf{X}$*. Assume that there exists a column or row that includes less than r sampled entries. Then, there exist infinitely many rank-r completion of the sampled matrix.*

Proof. We will prove the lemma when there exists such a column and the proof for the existence of such a row is similar. Assume that the first column includes less than r sampled entries. Note that there exists at least one completion (the original sampled matrix). Hence, there exists a basis $\mathbf{V} \in \mathbb{R}^{n\times r}$ such that each column of the sampled matrix (on the locations of the sampled entries) can be written as linear combination of the r columns of $\mathbf{V}$. In order to complete the proof, it suffices to show that there exist infinite many completions of the first column such that it is a linear combination of the columns of $\mathbf{V}$. Let $\mathbf{x}_1$ denote the first column of $\mathbf{X}$ and $\mathbf{v}_i$ denote the i-th column of $\mathbf{V}$, $i = 1, \ldots, r$. Then, we need to show that there exist infinitely many solutions to $(\mathbf{x}_1)_{\mathbf{\Omega}} = (a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r)_{\mathbf{\Omega}}$ in terms of the scalars a_i's, where $(\mathbf{x}_1)_{\mathbf{\Omega}}$ and $(a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r)_{\mathbf{\Omega}}$ represent the vectors consisting of those entries of $\mathbf{x}_1$ and $a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r$ that their corresponding locations in the first column of the sampled matrix have been sampled, respectively. Note that this system of equations includes less than r scalar equations and each equation is in terms of r variables.

The genericity assumption and existence of at least one completion simply conclude that there exist infinitely many solutions to $(\mathbf{x}_1)_{\mathbf{\Omega}} = (a_1\mathbf{v}_1 + \cdots + a_r\mathbf{v}_r)_{\mathbf{\Omega}}$. $\square$

Consider a matrix $\mathbf{U} \in \mathbb{R}^{4\times5}$, where $\mathbf{U} = [\mathbf{U}_1|\mathbf{U}_2]$, $\mathbf{U}_1 \in \mathbb{R}^{4\times2}$ (the first two columns) and $\mathbf{U}_2 \in \mathbb{R}^{4\times3}$ (the last three columns). Assume that $r_1 = 1$, $r_2 = 2$ and $r = 2$. Moreover, suppose that the sampled entries of $\mathbf{U}$ are shown below.

$$\underbrace{\begin{bmatrix} \overbrace{\times \quad \times}^{\mathbf{U}_1} & \overbrace{\times \quad \times \quad \times}^{\mathbf{U}_2} \\ \times \quad - & \times \quad \times \quad \times \\ \times \quad - & \times \quad - \quad - \\ - \quad - & - \quad \times \quad \times \end{bmatrix}}_{\mathbf{U}}$$

We have the following observations about the number of completions of each matrix.

- *Given* $r_1 = 1$, $\mathbf{U}_1$ *is infinitely completable*: For any value of the (4,1)-th entry of $\mathbf{U}_1$, there exists exactly one completion of $\mathbf{U}_1$. Hence, there exist infinitely completions of $\mathbf{U}_1$. We can also verify infinite completability via Lemma 1 since the fourth row of $\mathbf{U}_1$ has no sampled entry.
- *Given* $r_2 = 2$, $\mathbf{U}_2$ *is infinitely completable*: Observe that each value of the (3,2)-th entry of $\mathbf{U}_2$, corresponds to one completion of $\mathbf{U}_2$. As a result, there are infinitely many completions of $\mathbf{U}_2$. We can also verify infinite completability via Lemma 1 since the third row of $\mathbf{U}_2$ has only one sampled entry.
- *Given* $r = 2$, $\mathbf{U}$ *is infinitely completable*: Note that for any value of the (2,2)-th entry of $\mathbf{U}$, there exists at least one completion of $\mathbf{U}$ (as the second column of $\mathbf{U}$ is a linear combination of two vectors and only one entry of this column is known), and therefore $\mathbf{U}$ is infinitely completable. We can also verify infinite completability via Lemma 1 since the second column of $\mathbf{U}$ has only one sampled entry.
- *For almost every matrix* $\mathbf{U}$, *given* $r_1 = 1$, $r_2 = 2$ *and* $r = 2$, $\mathbf{U}$ *is finitely completable*: We prove this statement in Section Appendix A by applying Theorem 1 which takes advantage of a geometric analysis on the manifold structure for multi-view data (which is not Grassmannian manifold) to incorporate all three rank constraints simultaneously.

A larger example is provided in Section Appendix B to show the advantage of multi-view analysis when $\max\{r_1, r_2\} < r < r_1 + r_2$.

As mentioned earlier, the multi-view learning problem has many applications in various areas. Hence, such analysis (that outperforms applying the single-view analysis multiple times) characterizes the fundamental bounds for the corresponding applications.

We next construct a constraint matrix $\breve{\mathbf{\Omega}}$ based on $\mathbf{\Omega}$ such that each column of $\breve{\mathbf{\Omega}}$ represents exactly one of the polynomials in $\mathcal{P}(\mathbf{\Omega})$. Consider an arbitrary column of the first view $\mathbf{U}_1(:, i)$, where $i \in \{1, \dots, m_1\}$. Let $l_i = N_{\mathbf{\Omega}}(\mathbf{U}_1(:, i))$ denote the number of observed entries in the i-th column of the first view. Assumption 1 results that $l_i \geq r_1$.

We construct $l_i - r_1$ columns with binary entries based on the locations of the observed entries in $\mathbf{U}_1(:, i)$ such that each column has exactly $r_1 + 1$ entries equal to one. Assume that $x_1, \dots, x_{l_i}$ be the row indices of all observed entries in this column. Let $\mathbf{\Omega}_1^i$ be the corresponding $n \times (l_i - r_1)$ matrix to this column which is defined as the following: for any $j \in \{1, \dots, l_i - r_1\}$, the j-th column has the value 1 in rows $\{x_1, \dots, x_{r_1}, x_{r_1+j}\}$ and zeros elsewhere. Define the binary constraint matrix of the first view as $\breve{\mathbf{\Omega}}_1 = \left[\mathbf{\Omega}_1^1|\mathbf{\Omega}_1^2 \dots |\mathbf{\Omega}_1^{m_1}\right] \in \mathbb{R}^{n\times K_1}$ [34], where $K_1 = N_{\mathbf{\Omega}}(\mathbf{U}_1) - m_1 r_1$.

Similarly, we construct the binary constraint matrix $\breve{\mathbf{\Omega}}_2 \in \mathbb{R}^{n\times K_2}$ for the second view, where $K_2 = N_{\mathbf{\Omega}}(\mathbf{U}_2) - m_2 r_2$. Define the constraint matrix of $\mathbf{U}$ as $\breve{\mathbf{\Omega}} = [\breve{\mathbf{\Omega}}_1|\breve{\mathbf{\Omega}}_2] \in \mathbb{R}^{n\times(K_1+K_2)}$. For any subset of columns $\breve{\mathbf{\Omega}}'$ of $\breve{\mathbf{\Omega}}$, $\mathcal{P}(\breve{\mathbf{\Omega}}')$ denotes the subset of $\mathcal{P}(\mathbf{\Omega})$ that corresponds to $\breve{\mathbf{\Omega}}'$.

Consider an example, where matrix $\mathbf{U} = [\mathbf{U}_1|\mathbf{U}_2] \in \mathbb{R}^{4\times4}$ and $\mathbf{U}_1 \in \mathbb{R}^{4\times2}$ is the first view and $\mathbf{U}_2 \in \mathbb{R}^{4\times2}$ is the second view. The samples that are used to obtain $(\mathbf{T}_1, \mathbf{T}_2)$ are colored as red in the following. Assume that $r_1 = 1$, $r_2 = 2$ and $r = 2$. Then, the constraint matrix is as the following.

$$\mathbf{U} = \begin{bmatrix} \overbrace{\times \quad \times}^{\mathbf{U}_1} & \overbrace{\times \quad \times}^{\mathbf{U}_2} \\ \times \quad - & \times \quad \times \\ \times \quad - & \times \quad - \\ - \quad - & \times \quad \times \end{bmatrix}, \breve{\mathbf{\Omega}} = \begin{bmatrix} \overbrace{1 \quad 1}^{\breve{\mathbf{\Omega}}_1} & \overbrace{1 \quad 1 \quad 1}^{\breve{\mathbf{\Omega}}_2} \\ 1 \quad 0 & 1 \quad 1 \quad 1 \\ 0 \quad 1 & 1 \quad 0 \quad 0 \\ 0 \quad 0 & 0 \quad 1 \quad 1 \end{bmatrix}.$$

Assume that $\breve{\mathbf{\Omega}}'$ is an arbitrary subset of columns of the constraint matrix $\breve{\mathbf{\Omega}}$. Then, $\breve{\mathbf{\Omega}}_1'$ and $\breve{\mathbf{\Omega}}_2'$ denote the columns that correspond to the first and second views, respectively. Similarly, assume that $\mathbf{\Omega}'$ is an arbitrary subset of columns of $\mathbf{\Omega}$. Then, $\mathbf{\Omega}_1'$ and $\mathbf{\Omega}_2'$ denote the columns that correspond to the first view and second view, respectively. Moreover, for any matrix $\mathbf{X}$, $c(\mathbf{X})$ denotes the number of columns of $\mathbf{X}$ and $g(\mathbf{X})$ denotes the number of nonzero rows of $\mathbf{X}$. A submatrix of the constraint matrix is called a **proper** submatrix if its columns correspond to different columns of the sampling pattern.

We define $r_1' = r - r_2$, $r_2' = r - r_1$ and $r' = r - r_1' - r_2' = r_1 + r_2 - r$. Observe that $r_1 \leq r$, $r_2 \leq r$ and $r \leq r_1 + r_2$. Suppose that the basis $\mathbf{V} \in \mathbb{R}^{n\times r}$ is such that its first r_1 columns constitute a basis for the first view $\mathbf{U}_1$, its last r_2 columns constitute a basis for the second view $\mathbf{U}_2$, and all r columns of $\mathbf{V}$ constitute a basis for $\mathbf{U} = [\mathbf{U}_1|\mathbf{U}_2]$, as shown in Fig. 1. Assume that $\mathbf{V} = [\mathbf{V}_1|\mathbf{V}_2|\mathbf{V}_3]$, where $\mathbf{V}_1 \in \mathbb{R}^{n\times r_1'}$, $\mathbf{V}_2 \in \mathbb{R}^{n\times r'}$ and $\mathbf{V}_3 \in \mathbb{R}^{n\times r_2'}$. Then, $[\mathbf{V}_1|\mathbf{V}_2]$ is a basis for $\mathbf{U}_1$ and $[\mathbf{V}_2|\mathbf{V}_3]$ is a basis for $\mathbf{U}_2$. Hence, there exist $\mathbf{T}_1 \in \mathbb{R}^{r_1\times m_1}$ and $\mathbf{T}_2 \in \mathbb{R}^{r_2\times m_2}$ such that

$$\mathbf{U}_1 = [\mathbf{V}_1|\mathbf{V}_2] \cdot \mathbf{T}_1, \tag{1a}$$

$$\mathbf{U}_2 = [\mathbf{V}_2|\mathbf{V}_3] \cdot \mathbf{T}_2. \tag{1b}$$

For any $i_1 \in \{1, \dots, m_1\}$ and $i_2 \in \{1, \dots, m_2\}$, (1) can be written as

$$\mathbf{U}_1(:, i_1) = [\mathbf{V}_1|\mathbf{V}_2] \cdot \mathbf{T}_1(:, i_1), \tag{2a}$$

$$\mathbf{U}_2(:, i_2) = [\mathbf{V}_2|\mathbf{V}_3] \cdot \mathbf{T}_2(:, i_2). \tag{2b}$$

Let $\mathcal{M}(r, r_1, r_2, \mathbb{R}^n)$ denote the manifold structure of subspaces $\mathbf{V}$ described above for the multi-view matrix and define $\mathbb{P}_{\mathcal{M}}$ as the uniform measure on this manifold. Moreover, define $\mathbb{P}_L$ as the Lebesgue measure on $\mathbb{R}^{r\times(m_1+m_2)}$. We assume that $\mathbf{U}$ is chosen generically from $\mathcal{M}(r, r_1, r_2, \mathbb{R}^n)$, or in other words, the entries of $\mathbf{U}$ are drawn independently with respect to Lebesgue measure on $\mathcal{M}(r, r_1, r_2, \mathbb{R}^n)$. Hence, any statement that holds for $\mathbf{U}$, it also holds for almost every (with probability one) data of the same size and rank with respect to the product measure $\mathbb{P}_{\mathcal{M}} \times \mathbb{P}_L$. Note that according to Proposition 2, each multi-view data $\mathbf{U}$ can be uniquely represented in terms of a subspace $\mathbf{V} \in \mathcal{M}(r, r_1, r_2, \mathbb{R}^n)$.

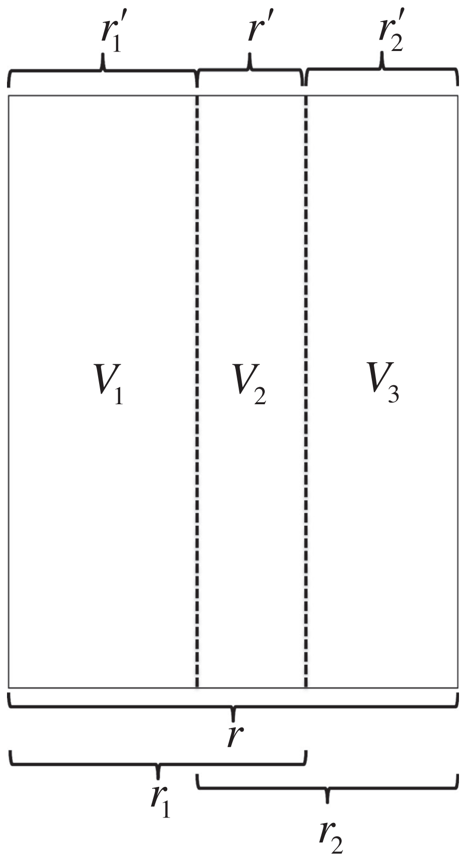


Fig. 1. A basis $\mathbf{V}$ for the sampled matrix $\mathbf{U}$.

Note that it can be concluded from Bernstein's theorem [44] that in a system of n polynomials in n variables with each consisting of a given set of monomials such that the coefficients are chosen with respect to the Lebesgue measure on $\mathcal{M}(r, r_1, r_2, \mathbb{R}^n)$, the n polynomials are algebraically independent with probability one with respect to the product measure $\mathbb{P}_{\mathcal{M}} \times \mathbb{P}_L$, and therefore there exist only finitely many solutions (all given probabilities in this paper are with respect to this product measure). However, in the structure of the polynomials in our model, the set of involved monomials are different for different set of polynomials, and therefore to ensure algebraically independency we need to have for any selected subset of the original n polynomials, the number of involved variables should be more than the number of selected polynomials.

Given all observed entries $\{\mathbf{U}(x_1, x_2) : \mathbf{\Omega}(x_1, x_2) = 1\}$, we are interested in finding the number of possible solutions in terms of entries of $(\mathbf{V}, \mathbf{T}_1, \mathbf{T}_2)$ (infinite, finite or unique) via investigating the algebraic independence among the polynomials. Throughout this paper, we make the following assumption.

Assumption 1. Any column of $\mathbf{U}_1$ includes at least r_1 observed entries and any column of $\mathbf{U}_2$ includes at least r_2 observed entries.

Observe that Assumption 1 leads to a total of at least $m_1 r_1 + m_2 r_2$ sampled entries of $\mathbf{U}$.

Lemma 2 ([45]). *Given a basis $\mathbf{V} = [\mathbf{V}_1 | \mathbf{V}_2 | \mathbf{V}_3]$ in (1), if Assumption 1 holds, then there exists a unique solution $(\mathbf{T}_1, \mathbf{T}_2)$ to the set of polynomials obtained from the sampled entries, with probability one. Moreover, if Assumption 1 does not hold, then there are infinite number of solutions $(\mathbf{T}_1, \mathbf{T}_2)$, with probability one.*

As a result of Lemma 2, Assumption 1 is necessary and sufficient for having finite (and unique) number of solutions to the completion problem if the basis is given.

Definition 1. Observe that given $\mathbf{V}$, each observed entry of $\mathbf{U}_1$ and $\mathbf{U}_2$ results in a degree-1 polynomial whose involved variables are the entries of the corresponding column of $\mathbf{T}_1$ and $\mathbf{T}_2$, respectively. We choose r_1 and r_2 observed entries of each column of $\mathbf{U}_1$ and $\mathbf{U}_2$, respectively, to obtain $(\mathbf{T}_1, \mathbf{T}_2)$. Let $\mathcal{P}(\mathbf{\Omega})$ denote all polynomials in terms of the entries of $\mathbf{V}$ obtained through the observed entries excluding the $m_1 r_1 + m_2 r_2$ polynomials that were used to obtain $(\mathbf{T}_1, \mathbf{T}_2)$. Note that since $(\mathbf{T}_1, \mathbf{T}_2)$ is already solved in terms of $\mathbf{V}$, each polynomial in $\mathcal{P}(\mathbf{\Omega})$ is in terms of elements of $\mathbf{V}$.

Consider two bases $\mathbf{V}$ and $\mathbf{V}'$ for the matrix $\mathbf{U}$ with the structure in (1). We say that $\mathbf{V}$ and $\mathbf{V}'$ span the same space if and only if: (i) the spans of the first r_1 columns of $\mathbf{V}$ and $\mathbf{V}'$ are the same, (ii) the spans of the last r_2 columns of $\mathbf{V}$ and $\mathbf{V}'$ are the same, (iii) the spans of all columns of $\mathbf{V}$ and $\mathbf{V}'$ are the same.

Therefore, $\mathbf{V}$ and $\mathbf{V}'$ span the same space if and only if: (i) each column of $\mathbf{V}_1$ is a linear combination of the columns of $[\mathbf{V}_1' | \mathbf{V}_2']$, (ii) each column of $\mathbf{V}_2$ is a linear combination of the columns of $\mathbf{V}_2'$, and (iii) each column of $\mathbf{V}_3$ is a linear combination of the columns of $[\mathbf{V}_2' | \mathbf{V}_3']$. The following equivalence class partitions all possible bases such that any two bases in a class span the same space, i.e., the above-mentioned properties (i), (ii) and (iii) hold.

Definition 2. Define an equivalence class for all bases $\mathbf{V} \in \mathbb{R}^{n \times r}$ of the sampled matrix $\mathbf{U}$ such that two bases $\mathbf{V}$ and $\mathbf{V}'$ belong to the same class if there exist full rank matrices $\mathbf{A}_1 \in \mathbb{R}^{r_1 \times r_1'}$, $\mathbf{A}_2 \in \mathbb{R}^{r' \times r'}$ and $\mathbf{A}_3 \in \mathbb{R}^{r_2 \times r_2'}$ such that

$$\mathbf{V}_1 = [\mathbf{V}_1' | \mathbf{V}_2'] \cdot \mathbf{A}_1, \tag{3a}$$

$$\mathbf{V}_2 = \mathbf{V}_2' \cdot \mathbf{A}_2, \tag{3b}$$

$$\mathbf{V}_3 = [\mathbf{V}_2' | \mathbf{V}_3'] \cdot \mathbf{A}_3, \tag{3c}$$

where $\mathbf{V} = [\mathbf{V}_1 | \mathbf{V}_2 | \mathbf{V}_3]$, $\mathbf{V}' = [\mathbf{V}_1' | \mathbf{V}_2' | \mathbf{V}_3']$, $\mathbf{V}_1, \mathbf{V}_1' \in \mathbb{R}^{n \times r_1'}$, $\mathbf{V}_2, \mathbf{V}_2' \in \mathbb{R}^{n \times r'}$ and $\mathbf{V}_3, \mathbf{V}_3' \in \mathbb{R}^{n \times r_2'}$.

Note that (3) leads to the fact that the dimension of the space of all bases $\mathbf{V}$ in one particular class, i.e., the degree of freedom for the bases in one particular class, is equal to $nr - r_1 r_1' - r' r' - r_2 r_2' = nr - r^2 - r_1^2 - r_2^2 + r(r_1 + r_2)$.

Definition 3. (Canonical basis) As shown in Fig. 2, denote

$$\mathbf{B}_1 = \mathbf{V}\left(1 : r_1', 1 : r_1'\right) \in \mathbb{R}^{r_1' \times r_1'}, \tag{4a}$$

$$\mathbf{B}_2 = \mathbf{V}\left(1 : r_2', 1 + r_1 : r_2' + r_1\right) \in \mathbb{R}^{r_2' \times r_2'}, \tag{4b}$$

$$\mathbf{B}_3 = \mathbf{V}\Big(1 + \max(r_1', r_2') : r' + \max(r_1', r_2'), \; 1 + r_1' : r' + r_1'\Big) \in \mathbb{R}^{r' \times r'}, \tag{4c}$$

$$\mathbf{B}_4 = \mathbf{V}\Big(1 + \max(r_1', r_2') : r' + \max(r_1', r_2'), \; 1 : r_1'\Big) \in \mathbb{R}^{r' \times r_1'}, \tag{4d}$$

$$\mathbf{B}_5 = \mathbf{V}\Big(1 + \max(r_1', r_2') : r' + \max(r_1', r_2'), \; 1 + r_1 : r_2' + r_1\Big) \in \mathbb{R}^{r' \times r_2'}. \tag{4e}$$

Then, we call $\mathbf{V}$ a canonical basis if $\mathbf{B}_1 = \mathbf{I}_{r_1'}$, $\mathbf{B}_2 = \mathbf{I}_{r_2'}$, $\mathbf{B}_3 = \mathbf{I}_{r'}$, $\mathbf{B}_4 = \mathbf{0}_{r' \times r_1'}$ and $\mathbf{B}_5 = \mathbf{0}_{r' \times r_2'}$.

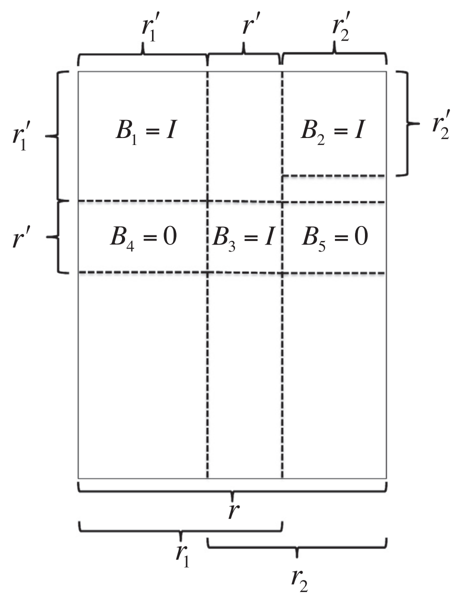


Fig. 2. A canonical basis.

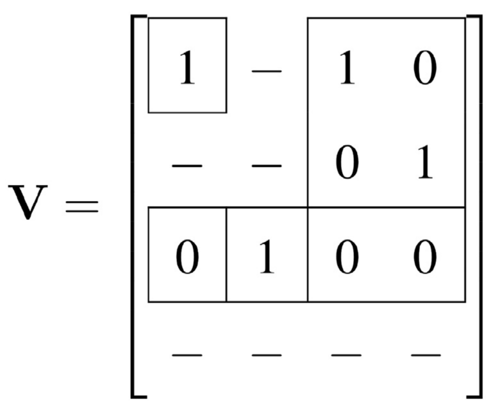


Fig. 3. The canonical basis for Example 1.

Example 1. Consider an example in which $\mathbf{U} = [\mathbf{U}_1|\mathbf{U}_2] \in \mathbb{R}^{4\times 7}$ and $\mathbf{U}_1 \in \mathbb{R}^{4\times 3}$ is the first view and $\mathbf{U}_2 \in \mathbb{R}^{4\times 4}$ is the second view. Assume that $r_1 = 2$, $r_2 = 3$ and $r = 4$. Then, the corresponding canonical basis is as shown in Fig. 3.

Observe that $r^2 + r_1^2 + r_2^2 - r(r_1 + r_2) = 9$ of the entries are known.

The following proposition shows the uniqueness of the canonical basis for the single-view matrix.

Proposition 1. *Assume that $\mathbf{X} \in \mathbb{R}^{n_1\times n_2}$ is generically chosen from the manifold of $n_1 \times n_2$ matrices of rank r. For almost every $\mathbf{X}$, there exists a unique basis $\mathbf{Y} \in \mathbb{R}^{n_1\times r}$ for $\mathbf{X}$ such that $\mathbf{Y}(1:r, 1:r) = \mathbf{I}_r$, where $\mathbf{Y}$ is a basis for $\mathbf{X}$ if each column of $\mathbf{X}$ can be written as a linear combination of the columns of $\mathbf{Y}$ and $\mathbf{Y}(1\colon r,\ 1\colon r)$ represents the submatrix of $\mathbf{Y}$ that consists of the first r columns and the first r rows and $\mathbf{I}_r$ denotes the $r \times r$ identity matrix.*

Proof. Consider an arbitrary basis $\mathbf{Y}' \in \mathbb{R}^{n_1\times r}$ for $\mathbf{X}$. Let $\mathcal{V}$ denote the set of all full rank $n_1 \times r$ matrices whose column span is equal to the column span of $\mathbf{Y}'$ and note that $\mathcal{V}$ is the set of all bases with r columns for $\mathbf{X}$. Consider an arbitrary member of set $\mathcal{V}$ and denote it by $\mathbf{Y}''$. Since, the column span of $\mathbf{Y}'$ and the column span of $\mathbf{Y}''$ are the same, each column of $\mathbf{Y}''$ can be written as a linear combination of columns of $\mathbf{Y}'$, and therefore there exists a unique full rank $\mathbf{Z} \in \mathbb{R}^{r\times r}$ such that $\mathbf{Y}'' = \mathbf{Y}'\mathbf{Z}$. Note that if $\mathbf{Z}$ is not full rank, we conclude $\mathbf{Y}''$ is not full rank as well, which contradicts the assumption.

Moreover, genericity of $\mathbf{X}$ results that each $r \times r$ submatrix of $\mathbf{Y}'$ is full rank, with probability one. This is because we have $\mathbf{X} = \mathbf{Y}'\mathbf{T}$ for some $\mathbf{T} \in \mathbb{R}^{r\times n_2}$, and therefore the fact that the submatrix consisting of any r rows of $\mathbf{X}$ is full rank results that the submatrix consisting of any r rows of $\mathbf{Y}'$ is full rank as well. Let $\mathbf{Z}_0$ denote the inverse of $\mathbf{Y}'(1\colon r,\ 1\colon r)$, i.e., $\mathbf{Z}_0 = \left(\mathbf{Y}'(1:r, 1:r)\right)^{-1}$. Therefore, $\mathbf{Y} = \mathbf{Y}'\mathbf{Z}_0$ is the unique basis for $\mathbf{X}$ that $\mathbf{Y}(1:r, 1:r) = \mathbf{I}_r$. □

The following proposition considers the multi-view data described in this paper and shows the uniqueness of the canonical basis.

Proposition 2. *For almost every $\mathbf{U}$, there exists a unique basis $\mathbf{V} \in \mathbb{R}^{n\times r}$ for $\mathbf{U}$ such that $\mathbf{V}$ satisfies the canonical pattern in Definition 3 and also its first r_1 columns constitute a basis for the first view $\mathbf{U}_1$, its last r_2 columns constitute a basis for the second view $\mathbf{U}_2$, and all r columns of $\mathbf{V}$ constitute a basis for $\mathbf{U} = [\mathbf{U}_1|\mathbf{U}_2]$.*

Proof. Consider an arbitrary basis $\mathbf{V}' = [\mathbf{V}_1'|\mathbf{V}_2'|\mathbf{V}_3']$ for $\mathbf{U}$ such that its first r_1 columns constitute a basis for the first view $\mathbf{U}_1$, its last r_2 columns constitute a basis for the second view $\mathbf{U}_2$, and all r columns of $\mathbf{V}$ constitute a basis for $\mathbf{U} = [\mathbf{U}_1|\mathbf{U}_2]$. Let $\mathcal{V}$ denote the set of all such bases for $\mathbf{U}$ and consider an arbitrary member of this set and denote it by $\mathbf{V} = [\mathbf{V}_1|\mathbf{V}_2|\mathbf{V}_3]$. Hence, according to the earlier discussion before Definition 2, Equs. (3a)–(3c) hold. This is because the column spans of the first r_1 columns of $\mathbf{V}$ and $\mathbf{V}'$, or the column spans of the last r_2 columns of $\mathbf{V}$ and $\mathbf{V}'$ and also the column spans of the all r columns of $\mathbf{V}$ and $\mathbf{V}'$ are the same.

Similar to the proof of Proposition 1, (3b) results that the unique $\mathbf{V}_2$ that can satisfy the pattern $\mathbf{B}_3 = \mathbf{I}_{r'}$ in Definition 3, can be obtained by $\mathbf{A}_2$ equal to the inverse of matrix $\mathbf{V}'\left(1 + \max(r_1', r_2') : r' + \max(r_1', r_2'), 1 + r_1' : r' + r_1'\right)$. In order to complete the proof, it suffices to show that there exists a unique $\mathbf{A}_1$ that results in satisfying the patterns $\mathbf{B}_1 = \mathbf{I}_{r_1'}$ and $\mathbf{B}_4 = \mathbf{0}_{r'\times r_1'}$ in Definition 3 for $\mathbf{V}_1$ (the uniqueness of $\mathbf{A}_3$ is similar to that for $\mathbf{A}_1$).

Let $\mathbf{A}_1' \in \mathbb{R}^{r_1\times r_1}$ denote the inverse of $r_1 \times r_1$ matrix $[\mathbf{V}_1|\mathbf{V}_2](1\colon r_1,\ 1\colon r_1)$. Note that existing of the inverse is a consequence of genericity assumption which results the submatrix consisting of any r_1 rows of $\mathbf{U}_1$ is full rank with probability one, and therefore the submatrix consisting of any r_1 rows of $[\mathbf{V}_1|\mathbf{V}_2]$ is full rank. Let $\mathbf{A}_1$ be the first r_1' columns of $\mathbf{A}_1'$. Then, $\mathbf{A}_1$ ensures that the patterns $\mathbf{B}_1 = \mathbf{I}_{r_1'}$ and $\mathbf{B}_4 = \mathbf{0}_{r'\times r_1'}$ in Definition 3 hold for $\mathbf{V}_1$. Finally, note that $\mathbf{A}_1$ is unique. Otherwise, there exist two different inverse matrices for the full rank $r_1 \times r_1$ matrix $[\mathbf{V}_1|\mathbf{V}_2](1\colon r_1,\ 1\colon r_1)$, which is contradiction. □

Remark 1. In order to prove there are finitely many completions for the matrix $\mathbf{U}$, it suffices to prove that given $\mathbf{T}_1$ and $\mathbf{T}_2$, there are finitely many canonical bases that fit in $\mathbf{U}$, where a basis fitting in $\mathbf{U}$ is equivalent to the existence of a completion of $\mathbf{U}$ such that each of its columns can be written as a linear combination of the corresponding basis.

Note that patterns $\mathbf{B}_1$ and $\mathbf{B}_4$ are in $\mathbf{V}_1$, patterns $\mathbf{B}_2$ and $\mathbf{B}_5$ are in $\mathbf{V}_3$, and pattern $\mathbf{B}_3$ is in $\mathbf{V}_2$. We can also easily show that any permutation of the rows of any of these patterns satisfies the property that in each class there exists exactly one basis with the permuted pattern.

4. Sampling guarantees for finite completability

In this section, we show that if the number of samples in each column satisfies a proposed lower bound, then the conditions stated in the statement of Theorem 1 on sampling pattern hold, i.e., $\mathbf{U}$ is finitely completable with high probability.

In order to prove this result, we will first find some results for the algerbraic independence of the polynomials. We first note that if Assumption 1 holds, then for almost every sampled matrix $\mathbf{U}$, there are at most finitely many bases that fit in $\mathbf{U}$ if and only if there exist $nr - r^2 - r_1^2 - r_2^2 + r(r_1 + r_2)$ algebraically independent polynomials in $\mathcal{P}(\mathbf{\Omega})$ (The detailed proof of this can be seen in [45]). The next result provides an upper bound on the maximum number of algebraically independent polynomials in any subset of columns of the constraint matrix $\breve{\mathbf{\Omega}}$.

Lemma 3. *Assume that Assumption 1 holds. Let $\breve{\mathbf{\Omega}}'$ be a proper subset of columns of the constraint matrix $\breve{\mathbf{\Omega}}$. Then, the maximum number of algebraically independent polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}}')$ is upper bounded by*

$$r_1'(g(\breve{\mathbf{\Omega}}_1') - r_1)^+ + r_2'(g(\breve{\mathbf{\Omega}}_2') - r_2)^+ + r'(g(\breve{\mathbf{\Omega}}') - r')^+. \quad (5)$$

Proof. The maximum number of algebraically independent polynomials in $\mathcal{P}(\breve{\mathbf{\Omega}}')$ is at most equal to the number of involved variables in the polynomials. Note that each observed entry of $\mathbf{U}_1$ results in a polynomial that involves all r_1 entries of a row of $\mathbf{V}_1$. As a result, the number of entries of $\mathbf{V}_1$ that are involved in the polynomials is exactly $r_1 g(\breve{\mathbf{\Omega}}_1')$. As mentioned earlier, the rows of patterns $\mathbf{B}_1$ and $\mathbf{B}_4$ in Definition 3 can be permuted such that exactly one basis in each class satisfies the new pattern. Hence, it can be permuted such that r_1 rows of $\mathbf{B}_1$ and $\mathbf{B}_4$ are a subset of the nonzero rows of $\breve{\mathbf{\Omega}}_1'$ since there are at least $r_1 + 1$ nonzero rows (any column of the constraint matrix of the first view includes exactly $r_1 + 1$ nonzero entries). Recall that the total number of known entries of $\mathbf{V}_1$ is the summation of the number of entries of $\mathbf{B}_1$ and $\mathbf{B}_4$, i.e., $r_1' r_1$. Therefore, the number of variables (unknown entries) of $\mathbf{V}_1$ that are involved in $\mathcal{P}(\breve{\mathbf{\Omega}}')$ is equal to $r_1'(g(\breve{\mathbf{\Omega}}_1') - r_1)^+$. Note that $g(\breve{\mathbf{\Omega}}_1') - r_1$ is negative if and only if $g(\breve{\mathbf{\Omega}}_1') = 0$.

Similarly, the number of unknown entries of $\mathbf{V}_2$ and $\mathbf{V}_3$ that are invloved in $\mathcal{P}(\breve{\mathbf{\Omega}}')$ are $r'(g(\breve{\mathbf{\Omega}}') - r')^+$ and $r_2'(g(\breve{\mathbf{\Omega}}_2') - r_2)^+$, respectively. Therefore, the number of unknown entries of basis $\mathbf{V}$ that are involved in $\mathcal{P}(\breve{\mathbf{\Omega}}')$ is equal to $r_1'(g(\breve{\mathbf{\Omega}}_1') - r_1)^+ + r_2'(g(\breve{\mathbf{\Omega}}_2') - r_2)^+ + r'(g(\breve{\mathbf{\Omega}}') - r')^+$. □

A set of polynomials is called minimally algebraically dependent if the polynomials in that set are algebraically dependent but the polynomials in any of its proper subsets are algebraically independent. The next theorem gives the condition on $\breve{\mathbf{\Omega}}$ that is equivalent with existence of $nr - r^2 - r_1^2 - r_2^2 + r(r_1 + r_2)$ algebraically independent polynomials in $\mathcal{P}(\mathbf{\Omega})$,thereby providing the condition on $\breve{\mathbf{\Omega}}$ for finite completability of $\mathbf{U}$.

Theorem 1. *Assume that Assumption 1 holds. For almost every $\mathbf{U}$, the sampled matrix $\mathbf{U}$ is finite completable if and only if there exists a proper subset of columns $\breve{\mathbf{\Omega}}' \in \mathbb{R}^{n \times m}$ of the constraint matrix $\breve{\mathbf{\Omega}}$ such that $m = nr - r^2 - r_1^2 - r_2^2 + r(r_1 + r_2)$ and for any subset of columns $\breve{\mathbf{\Omega}}''$ of $\breve{\mathbf{\Omega}}'$ the following inequality holds*

$$r_1'(g(\breve{\mathbf{\Omega}}_1'') - r_1)^+ + r_2'(g(\breve{\mathbf{\Omega}}_2'') - r_2)^+ + r'(g(\breve{\mathbf{\Omega}}'') - r')^+ \geq c(\breve{\mathbf{\Omega}}''). \quad (6)$$

Proof. The result is an extension of the deterministic guarantee result in [37], using Lemma 3. The detailed proof is thus omitted. □

The proposed deterministic analysis and the condition on the sampling pattern in Theorem 1 have a combinatorial nature such that it takes exponential amount of time (in terms of the parameters of the problem) to verify the these geometric conditions on the sampling pattern. This combinatorial nature is consequence of NP-hardness of the problem (otherwise the problem would be polynomial solvable). Therefore, we are motivated to find a lower bound on the sampling pattern such that if that inequality holds, then we can guarantee these difficultly verifiable conditions hold with high probability. Hence, such analysis would have a very practical value as well. To this end, we propose a combinatorial analysis to find the required number of samples per column (under uniform sampling) such that the condition on the sampling pattern in Theorem 1 holds.

In order to find such bound on the sampling rate, we first provide the next lemma that will be used to prove Theorem 2. More specifically, in Theorem 2 we consider three disjoint sets of columns of $\mathbf{U}$ and apply Lemma 4 to each of them. Then, we combine the three sets of columns and show that they satisfy the conditions stated in the statement of Theorem 1. This lemma connects the assumption of having a certain number of samples per column (under uniform sampling) with a geometrical property on the location of the sampled entries, i.e., sampling pattern.

Lemma 4. *Assume that $r'' \leq \frac{n}{6}$ and also each column of $\mathbf{\Omega}$ includes at least l nonzero entries, where*

$$l > \max\left\{9 \ \log\left(\frac{n}{\epsilon}\right) + 3 \ \log\left(\frac{k}{\epsilon}\right) + 6, 2r''\right\}. \quad (7)$$

Let $\mathbf{\Omega}'$ be an arbitrary set of $n - r''$ columns of $\mathbf{\Omega}$. Then, with probability at least $1 - \frac{\epsilon}{k}$, every subset $\mathbf{\Omega}''$ of columns of $\mathbf{\Omega}'$ satisfies

$$g(\mathbf{\Omega}'') - r'' \geq c(\mathbf{\Omega}''). \quad (8)$$

Proof. Please refer to the proof of [34, Lemma 9]. Note that the only difference is that the last inequalities of (16) and (18) in [34] should now be upper bounded by $\frac{\epsilon}{rd}$ instead of $\frac{\epsilon}{d^2}$. □

Theorem 2. *Assume that the following inequalities hold*

$$\frac{n}{6} \geq \max\{r_1, r_2, r'\}, \quad (9)$$

$$m_1 \geq r_1'(n - r_1), \quad (10)$$

$$m_2 \geq r_2'(n - r_2), \quad (11)$$

$$m_1 + m_2 \geq r_1'(n - r_1) + r_2'(n - r_2) + r'(n - r'). \quad (12)$$

Moreover, assume that each column of $\mathbf{\Omega}$ includes at least l nonzero entries, where

$$l > \max\left\{9 \ \log\left(\frac{n}{\epsilon}\right) + 3 \ \max\left\{\log\left(\frac{3r_1'}{\epsilon}\right), \log\left(\frac{3r_2'}{\epsilon}\right), \log\left(\frac{3r'}{\epsilon}\right)\right\} + 6, 2r_1, 2r_2\right\}. \quad (13)$$

Then, the multi-view data $\mathbf{U}$ has only finitely many possible completions with probability at least $1 - \epsilon$.

Proof. Let $\mathbf{\Omega}_1'$ be an arbitrary set of $n - r_1$ columns of $\mathbf{\Omega}_1$. Note that having (13), it is easy to see that (7) holds with k and r'' replaced by $3r_1'$ and r_1, respectively. Hence, having (9), Lemma 4 results that any subset of columns $\mathbf{\Omega}_1''$ of $\mathbf{\Omega}_1'$ satisfies

$$g(\mathbf{\Omega}_1'') - r_1 \geq c(\mathbf{\Omega}_1''), \quad (14)$$

with probability at least $1 - \frac{\epsilon}{3r_1'}$. According to Lemma 5 below and by setting $r = r_1$, as a subset of columns $\mathbf{\Omega}_1'$ of $\mathbf{\Omega}_1$ satisfies (14), there exists a subset of columns $\breve{\mathbf{\Omega}}_1'$ of the constraint matrix of the first view $\breve{\mathbf{\Omega}}_1$ (corresponding columns to the columns of $\mathbf{\Omega}_1'$) that satisfies (14) as well.

Assumption (10) results that $\boldsymbol{\Omega}_1$ includes at least $r'_1(n-r_1)$ columns or in other words, r'_1 disjoint sets of columns each including $n-r_1$ columns. All r'_1 disjoint sets satisfy property (i) simultaneously with probability at least $1-\frac{\epsilon}{3}$. Therefore, there exist r'_1 disjoint sets of columns each including $n-r_1$ columns of the constraint matrix of the first view $\breve{\boldsymbol{\Omega}}_1$, and also all r'_1 disjoint sets satisfy (14), simultaneously with probability at least $1-\frac{\epsilon}{3}$. Let $\breve{\breve{\boldsymbol{\Omega}}}_1$ denote the union of the r'_1 mentioned sets of columns.

Consider any subset of columns $\breve{\breve{\boldsymbol{\Omega}}}'_1$ of $\breve{\breve{\boldsymbol{\Omega}}}_1$ and define $\breve{\breve{\boldsymbol{\Omega}}}'_{1,i}$ as the intersection of $\breve{\breve{\boldsymbol{\Omega}}}'_1$ and the i-th set among the mentioned r'_1 sets for $i=1,\dots,r'_1$. Without loss of generality, assume that $\max_{1\le i\le r'_1}\{c(\breve{\breve{\boldsymbol{\Omega}}}'_{1,i})\} = c(\breve{\breve{\boldsymbol{\Omega}}}'_{1,1})$. Then,

$$c(\breve{\breve{\boldsymbol{\Omega}}}'_1) = \sum_{i=1}^{r'_1} c(\breve{\breve{\boldsymbol{\Omega}}}'_{1,i}) \le r'_1 c(\breve{\breve{\boldsymbol{\Omega}}}'_{1,i}) \le$$
$$r'_1(g(\breve{\breve{\boldsymbol{\Omega}}}'_{1,1}) - r_1)^+ \le r'_1(g(\breve{\breve{\boldsymbol{\Omega}}}'_1) - r_1)^+, \tag{15}$$

where the second inequality follows from (14). Therefore, we have

$$c(\breve{\breve{\boldsymbol{\Omega}}}'_1) \le r'_1(g(\breve{\breve{\boldsymbol{\Omega}}}'_1) - r_1)^+. \tag{16}$$

Note that having (13), it is easy to see that (7) holds with k and r'' replaced by $3r'_2$ and r_2, respectively. Moreover, recall that $r' = r_1 + r_2 - r \le \min\{r_1, r_2\}$, and therefore, having (13), it is easy to see that (7) holds with k and r'' replaced by $3r'$ and r', respectively. As a result, similarly, having (9) and (11), $\breve{\breve{\boldsymbol{\Omega}}}_2$ includes $r'_2(n-r_2)$ columns $\breve{\breve{\boldsymbol{\Omega}}}_2$ that with probability at least $1-\frac{\epsilon}{3}$ for any subset of it $\breve{\breve{\boldsymbol{\Omega}}}'_2$ we have

$$c(\breve{\breve{\boldsymbol{\Omega}}}'_2) \le r'_2(g(\breve{\breve{\boldsymbol{\Omega}}}'_2) - r_2)^+. \tag{17}$$

Using (12), $\boldsymbol{\Omega}$ includes $r'(n-r')$ columns $\bar{\boldsymbol{\Omega}}$ (disjoint from $\bar{\boldsymbol{\Omega}}_1$ and $\bar{\boldsymbol{\Omega}}_2$ corresponding to $\breve{\breve{\boldsymbol{\Omega}}}_1$ and $\breve{\breve{\boldsymbol{\Omega}}}_2$). Similar to $\breve{\breve{\boldsymbol{\Omega}}}_1$ and $\breve{\breve{\boldsymbol{\Omega}}}_2$, $\breve{\boldsymbol{\Omega}}$ includes $r'(n-r')$ columns $\breve{\breve{\boldsymbol{\Omega}}}$ (disjoint from $\breve{\breve{\boldsymbol{\Omega}}}_1$ and $\breve{\breve{\boldsymbol{\Omega}}}_2$) that with probability at least $1-\frac{\epsilon}{3}$ for any subset of columns of it $\breve{\breve{\boldsymbol{\Omega}}}'$ we have

$$c(\breve{\breve{\boldsymbol{\Omega}}}') \le r'(g(\breve{\breve{\boldsymbol{\Omega}}}') - r')^+. \tag{18}$$

Therefore, any subset of columns of $\breve{\breve{\boldsymbol{\Omega}}}_1$ satisfies (16) and any subset of $\breve{\breve{\boldsymbol{\Omega}}}_2$ satisfies (17) and any subset of $\breve{\breve{\boldsymbol{\Omega}}}$ satisfies (18) simultaneously with probability at least $1-\epsilon$. Define $\breve{\boldsymbol{\Omega}}' = [\breve{\breve{\boldsymbol{\Omega}}}_1 | \breve{\breve{\boldsymbol{\Omega}}}_2 | \breve{\breve{\boldsymbol{\Omega}}}] \in \mathbb{R}^{n\times m}$, where

$$m = r'(n-r') + r'_1(n-r_1) + r'_2(n-r_2) =$$
$$nr - r^2 - r_1^2 - r_2^2 + r(r_1 + r_2). \tag{19}$$

Let $\breve{\boldsymbol{\Omega}}''$ be a subset of columns of $\breve{\boldsymbol{\Omega}}'$ and define $\breve{\boldsymbol{\Omega}}''_1$, $\breve{\boldsymbol{\Omega}}''_2$ and $\breve{\boldsymbol{\Omega}}''_3$ as the intersection of $\breve{\boldsymbol{\Omega}}'$ with $\breve{\breve{\boldsymbol{\Omega}}}_1$, $\breve{\breve{\boldsymbol{\Omega}}}_2$ and $\breve{\breve{\boldsymbol{\Omega}}}$, respectively. Consequently, with probability at least $1-\epsilon$

$$c(\breve{\boldsymbol{\Omega}}'') = \sum_{i=1}^{3} c(\breve{\boldsymbol{\Omega}}''_i) \le r'_1(g(\breve{\boldsymbol{\Omega}}''_1) - r_1)^+ +$$
$$r'_2(g(\breve{\boldsymbol{\Omega}}''_2) - r_2)^+ + r'(g(\breve{\boldsymbol{\Omega}}''_3) - r')^+, \tag{20}$$

and therefore according to Theorem 1, $\mathbf{U}$ is finite completable with probability at least $1-\epsilon$. □

The following lemma is taken from [37, Lemma 8]. The lemma connects the sampling pattern and the constraint matrix (since we eventually need to verify the geometry pattern given in Theorem 1) by showing the equivalency of a geometrical property on the sampling pattern and a similar geometrical property on the constraint matrix.

Lemma 5. *Let R be a given nonnegative integer. Assume that there exists a matrix $\boldsymbol{\Omega}'$ such that it consists of $n-R$ columns of $\boldsymbol{\Omega}$ and each column of $\boldsymbol{\Omega}'$ includes at least $R+1$ nonzero entries and satisfies the following property:*

- *Denote an arbitrary matrix obtained by choosing any subset of the columns of $\boldsymbol{\Omega}'$ by $\boldsymbol{\Omega}''$. Then,*

$$g(\boldsymbol{\Omega}'') - R \ge c(\boldsymbol{\Omega}''). \tag{21}$$

Then, there exists a matrix $\breve{\boldsymbol{\Omega}}'$ with the same size as $\boldsymbol{\Omega}'$ such that: each column has exactly $R+1$ entries equal to one, and if $\breve{\boldsymbol{\Omega}}'(x,y) = 1$ then we have $\boldsymbol{\Omega}'(x,y) = 1$. Moreover, $\breve{\boldsymbol{\Omega}}'$ satisfies the above-mentioned property.

The next result finds a condition on the sampling probability that results (13).

Lemma 6. *Assume that the inequalities (9)–(12) hold. Moreover, assume that each entry of $\mathbf{U}$ is independently observed with probability p, where*

$$p > \frac{1}{n}\max\left\{9\ \log\left(\frac{n}{\epsilon}\right) + 3\ \max\left\{\log\left(\frac{3r'_1}{\epsilon}\right),\right.\right.$$
$$\left.\left.\log\left(\frac{3r'_2}{\epsilon}\right), \log\left(\frac{3r'}{\epsilon}\right)\right\} + 6, 2r_1, 2r_2\right\} + \frac{1}{\sqrt[4]{n}}. \tag{22}$$

Then, with probability at least $(1-\epsilon)\left(1-\exp(-\frac{\sqrt{n}}{2})\right)^{m_1+m_2}$, $\mathbf{U}$ is finitely completable.

Proof. Consider a vector with n entries where each entry is observed with probability p independently from the other entries. The authors of [37] showed that for $p > p' = \frac{k}{n} + \frac{1}{\sqrt[4]{n}}$, more than k entries are observed with probability at least $\left(1-\exp(-\frac{\sqrt{n}}{2})\right)$. Using this result, we note that the number of observed entries of each of the m_1+m_2 columns satisfies (13) with probability at least $\left(1-\exp(-\frac{\sqrt{n}}{2})\right)$. Hence, the proof follows using Theorem 2. □

5. Sampling guarantees for unique completability

Theorem 1 gives the necessary and sufficient condition on sampling pattern for finite completability. Hence, even one sample short of the condition in Theorem 1 results in infinite number of completions with probability one. We first provide an example for single-view matrix with exactly two completions to emphasize that finite completability does not necessarily result in unique completability (we can easily extend the example to multi-view matrix).

Example 2. Assume that the sampled matrix $\mathbf{U} \in \mathbb{R}^{5\times 4}$ is given as the incomplete matrix shown in Fig. 4.

Moreover, assume that $\text{rank}(\mathbf{U}) = 2$. In [37], it is shown that there exist exactly two completions of $\mathbf{U}$ as given by the two complete matrices above.

We show that adding a mild condition to the conditions obtained in the analysis for Problem (i) leads to unique completability. To this end, we obtain multiple sets of minimally algebraically dependent polynomials and show that the variables involved in these polynomials can be determined uniquely, and therefore entries of $\mathbf{U}$ can be determined uniquely.

Recall that there exists at least one completion of $\mathbf{U}$ since the original multi-view matrix that is sampled satisfies the rank constraints. The following lemma is a re-statement of Lemma 25 in [38]. Note that this lemma is also an adaptation of Lemma 7 in [34] or Theorem 1 in [46].

1		1	$-\frac{1}{2}$
-4	2	-1	
0	1		2
1		4	
	4	-2	$\frac{3}{2}$

$\Rightarrow$

1	$-\frac{21}{32}$	1	$-\frac{1}{2}$
-4	2	-1	$\frac{3}{4}$
0	1	$-\frac{24}{5}$	2
1	$-\frac{41}{32}$	4	$-\frac{7}{4}$
-8	4	-2	$\frac{3}{2}$

1	-2	1	$-\frac{1}{2}$
-4	2	-1	-10
0	1	$-\frac{1}{2}$	2
1	-8	4	$-\frac{25}{2}$
$-\frac{39}{21}$	4	-2	$\frac{3}{2}$

Fig. 4. A matrix with exactly two completions.

Lemma 7. *Assume that Assumption 1 holds. Let $\breve{\boldsymbol{\Omega}}'$ be an arbitrary subset of columns of the constraint matrix $\breve{\boldsymbol{\Omega}}$. Assume that polynomials in $\mathcal{P}(\breve{\boldsymbol{\Omega}}')$ are minimally algebraically dependent. Then, all variables (unknown entries) of $\mathbf{V}$ that are involved in $\mathcal{P}(\breve{\boldsymbol{\Omega}}')$ can be determined uniquely.*

Theorem 3 below gives a sufficient conditions on sampling pattern for unique completability. To be more specific, condition (i) in the statement of Theorem 3, i.e., $nr - r^2 - r_1^2 - r_2^2 + r(r_1 + r_2)$ algebraically independent polynomials in terms of the entries of $\mathbf{V}$, results in finite completability. Hence, adding any single polynomial to them results in a set of algebraically dependent polynomials and using Lemma 7 some of the entries of basis $\mathbf{V}$ can be determined uniquely. Then, conditions (ii) and (iii) result in more polynomials such that all entries of $\mathbf{V}$ can be determined uniquely.

Theorem 3. *Suppose that Assumption 1 holds. Moreover assume that there exist disjoint proper subsets of columns $\breve{\boldsymbol{\Omega}}' \in \mathbb{R}^{n\times m}$, $\breve{\boldsymbol{\Omega}}'_1 \in \mathbb{R}^{n\times m'}$ and $\breve{\boldsymbol{\Omega}}'_2 \in \mathbb{R}^{n\times m''}$ of the constraint matrix $\breve{\boldsymbol{\Omega}}$ such that the following properties hold*

(i) *$m = nr - r^2 - r_1^2 - r_2^2 + r(r_1 + r_2)$ and for any subset of columns $\breve{\boldsymbol{\Omega}}''$ of the matrix $\breve{\boldsymbol{\Omega}}'$, (6) holds.*
(ii) *$\breve{\boldsymbol{\Omega}}'_1$ is a subset of columns of $\breve{\boldsymbol{\Omega}}_1$ (constraint matrix of the first view), $m' = n - r_1$ and for any subset of columns $\breve{\boldsymbol{\Omega}}''_1$ of the matrix $\breve{\boldsymbol{\Omega}}'_1$*

$$g(\breve{\boldsymbol{\Omega}}''_1) - r_1 \geq c(\breve{\boldsymbol{\Omega}}''_1). \tag{23}$$

(iii) *$\breve{\boldsymbol{\Omega}}'_2$ is a subset of columns of $\breve{\boldsymbol{\Omega}}_2$ (constraint matrix of the second view), $m'' = n - r_2$ and for any subset of columns $\breve{\boldsymbol{\Omega}}''_2$ of the matrix $\breve{\boldsymbol{\Omega}}'_2$*

$$g(\breve{\boldsymbol{\Omega}}''_2) - r_2 \geq c(\breve{\boldsymbol{\Omega}}''_2). \tag{24}$$

Then, there exists only one completion of $\mathbf{U}$ that satisfies all the three rank constraints with probability one.

Proof. According to Theorem 1, property (i) results that there are only finitely many completions of $\mathbf{U}$ that satisfy the rank constraints. We show that having properties (ii) and (iii) results in obtaining all entries of the basis uniquely, and therefore there exists only one completion of $\mathbf{U}$. According to Theorem 1, the $nr - r^2 - r_1^2 - r_2^2 + r(r_1 + r_2)$ polynomials in $\mathcal{P}(\breve{\boldsymbol{\Omega}}')$ are algebraically independent. As a result, by adding any single polynomial to this set, we will have a set of algebraically dependent polynomials.

Consider a single polynomial from $\mathcal{P}(\breve{\boldsymbol{\Omega}}'_1) \cup \mathcal{P}(\breve{\boldsymbol{\Omega}}'_2)$ and denote it by p_0. Hence, polynomials in set $p_0 \cup \mathcal{P}(\breve{\boldsymbol{\Omega}}')$ are algebraically dependent, and therefore there exists $\mathcal{P}'(p_0) \subseteq \{p_0 \cup \mathcal{P}(\breve{\boldsymbol{\Omega}}')\}$ such that $p_0 \in \mathcal{P}'(p_0)$ and polynomials in $\mathcal{P}'(p_0)$ are minimally algebraically dependent. Lemma 7 results that all variables involved in polynomials in $\mathcal{P}'(p_0)$ can be determined uniquely. The number entries of $\mathbf{V}$ that are involved in $\mathcal{P}'(p_0)$ is at least r_1 if $p_0 \in \mathcal{P}(\breve{\boldsymbol{\Omega}}'_1)$ and r_2 if $p_0 \in \mathcal{P}(\breve{\boldsymbol{\Omega}}'_2)$. This is because the number of entries of $\mathbf{V}$ that are involved in polynomials in $\mathcal{P}'(p_0)$ is at least equal to the number of entries of $\mathbf{V}$ that are involved in p_0. Hence, $\mathcal{P}'(p_0)$ results in r_1 or r_2 polynomials that each has a unique solution.

Similarly, consider any other polynomial p_1 in $\mathcal{P}(\breve{\boldsymbol{\Omega}}'_1) \cup \mathcal{P}(\breve{\boldsymbol{\Omega}}'_2)$ and note that polynomials in set $p_1 \cup \mathcal{P}(\breve{\boldsymbol{\Omega}}')$ are algebraically dependent. Hence, we can repeat the above procedure for p_0 for polynomial p_1. Repeating this procedure for any subset of polynomials in $\mathcal{P}(\breve{\boldsymbol{\Omega}}''_1) \cup \mathcal{P}(\breve{\boldsymbol{\Omega}}''_2) \subseteq \mathcal{P}(\breve{\boldsymbol{\Omega}}'_1) \cup \mathcal{P}(\breve{\boldsymbol{\Omega}}'_2)$ results in $r'_1(g(\breve{\boldsymbol{\Omega}}''_1) - r_1)^+ + r'_2(g(\breve{\boldsymbol{\Omega}}''_2) - r_2)^+ + r'(g(\breve{\boldsymbol{\Omega}}''_3) - r')^+$ polynomials (as this is the number of unknown entries involved in the polynomials $\mathcal{P}(\breve{\mathbf{\Omega}}'_1) \cup \mathcal{P}(\breve{\mathbf{\Omega}}'_2))$ and observe that (23) and (24) result that the number of involved unknown entries of basis is not less than the number of polynomials, and therefore they are independent. Moreover, observe that $\breve{\mathbf{\Omega}}'_1$ and $\breve{\mathbf{\Omega}}'_2$ are such that polynomials obtained via this procedure cover all entries of basis. Therefore, all entries of basis can be determined uniquely with probability one. □

The next theorem gives a probabilistic guarantee for satisfying the conditions in the statement of Theorem 3 or in other words, a probabilistic guarantee for unique completability. However, similar to Theorem 2, the condition on sampling pattern is in terms of the number of samples per column instead of the complicated conditions in the statement of Theorem 3 on the structure of sampling pattern.

Theorem 4. *Assume that the following inequalities hold*

$$\frac{n}{6} \geq \max\{r_1, r_2, r'\}, \tag{25}$$

$$m_1 \geq (r'_1 + 1)(n - r_1), \tag{26}$$

$$m_2 \geq (r'_2 + 1)(n - r_2), \tag{27}$$

$$m_1 + m_2 \geq (r'_1 + 1)(n - r_1) + (r'_2 + 1)(n - r_2) + r'(n - r'). \tag{28}$$

Moreover, assume that each column of $\mathbf{\Omega}$ *includes at least* l *nonzero entries, where*

$$l > \max\left\{9 \ \log\left(\frac{n}{\epsilon}\right) + 3 \ \max\left\{\log\left(\frac{6r'_1}{\epsilon}\right), \log\left(\frac{6r'_2}{\epsilon}\right), \log\left(\frac{6r'}{\epsilon}\right)\right\} + 6, 2r_1, 2r_2\right\}. \tag{29}$$

Then, with probability at least $1 - \epsilon$*, there exists exactly one completion of* $\mathbf{U}$*.*

Proof. According to the proof of Theorem 2, (29) results that there exists a subset of columns $\breve{\mathbf{\Omega}}' \in \mathbb{R}^{n \times m}$ of the constraint matrix $\breve{\mathbf{\Omega}}$ such that condition (i) in the statement of Theorem 3 is satisfied, with probability at least $1 - \frac{\epsilon}{2}$. Then, assumptions (26), (27) and (28) result that there exist $n - r_1$ columns $\breve{\mathbf{\Omega}}'_1$ of $\breve{\mathbf{\Omega}}_1$ and $n - r_2$ columns $\breve{\mathbf{\Omega}}'_2$ of $\breve{\mathbf{\Omega}}_2$ that are disjoint from $\breve{\mathbf{\Omega}}'$. This is easily verified by comparing assumptions (26), (27) and assumptions (10), (11) in Theorem 2.

Note that according to Lemma 4, (29) results that $\breve{\mathbf{\Omega}}'_1$ satisfies condition (ii) in the statement of Theorem 3 with probability at least $1 - \frac{\epsilon}{6}$. Similarly, (29) results that $\breve{\mathbf{\Omega}}'_2$ satisfies condition (iii) in the statement of Theorem 3 with probability at least $1 - \frac{\epsilon}{6}$. Therefore, all conditions in the statement of Theorem 3 are satisfied simultaneously with probability at least $1 - \frac{\epsilon}{2} - \frac{\epsilon}{6} - \frac{\epsilon}{6}$. Hence, according to Theorem 3, there exists only one completion of $\mathbf{U}$ with probability at least $1 - \epsilon$. □

Remark 2. Comparing assumptions (10)–(12) for finite completability with assumptions (26)–(28) for unique completability, we see there is a mild change, i.e., r_i for finiteness is replaced by $r_i + 1$ for uniqueness.

Moreover, the lower bound on the number of samples per column increases mildly from (13) for finiteness to (29) for uniqueness, i.e., the factor 3 in the log terms in (13) become 6 in (29).

Lemma 8. *Assume that the inequalities (25)–(28) hold. Moreover, assume that each entry of* $\mathbf{U}$ *is independently observed with probability* p*, where*

$$p > \frac{1}{n} \max\left\{9 \ \log\left(\frac{n}{\epsilon}\right) + 3 \ \max\left\{\log\left(\frac{6r'_1}{\epsilon}\right), \log\left(\frac{6r'_2}{\epsilon}\right), \log\left(\frac{6r'}{\epsilon}\right)\right\} + 6, 2r_1, 2r_2\right\} + \frac{1}{\sqrt[4]{n}}. \tag{30}$$

Then, with probability at least $(1 - \epsilon)\left(1 - \exp(-\frac{\sqrt{n}}{2})\right)^{m_1 + m_2}$, $\mathbf{U}$ *is uniquely completable.*

Proof. Using similar arguments as in the proof of Theorem 6, the number of observed entries of each of the $m_1 + m_2$ columns satisfies (29) with probability at least $\left(1 - \exp(-\frac{\sqrt{n}}{2})\right)$. Hence, the proof follows using Theorem 4. □

6. Numerical comparisons

As we mentioned earlier, the existent matrix analysis can be applied multiple times for each of the rank constraints individually and provide some weak condition for finite and unique completability of the data. However, one purpose of this is to provide a stronger and more efficient way for finding such conditions using an analysis on the manifold corresponding to all the rank constraints together. Note that the numerical performance is only one purpose since also the methodology of handling multiple rank constraints for deterministic analysis is one of the main purposes of this work.

Here we compare the lower bound on the number of samples per column obtained by the proposed analysis in this paper with the bound obtained by the method in [34]. Recall that the existing method on Grassmannian manifold in [34] provides a bound on the number of samples for finite completability for a matrix $\mathbf{U}$ given $\text{rank}(\mathbf{U}) = r$. Note that we can not use the analysis in [34] for our multi-view data structure unless we obtain the bound in [34] corresponding to $\mathbf{U}$, $\mathbf{U}_1$ and $\mathbf{U}_2$ respectively (a trivial approach of using the analysis in [34] in our problem) and then take the maximum of them, it results in the following bound on the number of samples for finite completability

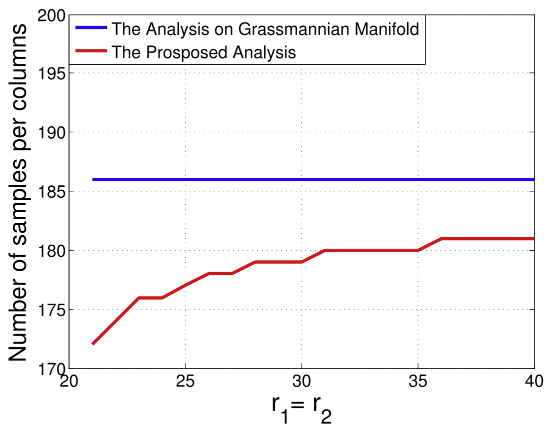
$$l > \max\left\{12 \ \log\left(\frac{n}{\epsilon}\right), 2r_1, 2r_2, 2r\right\}. \tag{31}$$

We consider a sampled data $\mathbf{U} = [\mathbf{U}_1|\mathbf{U}_2] \in \mathbb{R}^{500 \times 100000}$, where $\mathbf{U}_1, \mathbf{U}_2 \in \mathbb{R}^{500 \times 50000}$, i.e., $n = 500$ and $m_1 = m_2 = 50000$. In Fig. 5 we plot the bounds given in (13) for finite completability and compare it with the one in (31), as a function of the value $r_1 = r_2$, for $r = 40$, $r = 60$ and $r = 100$, with $\epsilon = 0.0001$. Recall that $r_1, r_2 \leq r$ and $r \leq r_1 + r_2$. It is seen that our proposed method requires less number of samples per column compared with the method in [34]. Note that given the large number of columns, i.e., $m = m_1 + m_2 = 10^5$, this leads to significantly less amount of sampled data.

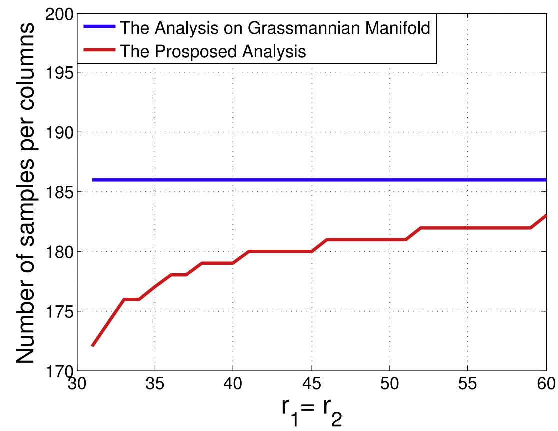
Note that the curves are not continuous as we need to apply the ceiling operator to the non-integer numbers in (13) and (31). Moreover, note that as both bounds in (13) and (31) are equal to the maximum of two terms: (i) one is on the order of $\log(n)$ or $\log(n) + \log(r)$, and (ii) one is linear in r. Hence, by increasing the value of r, eventually it will be a linear function of r, as seen in Fig. 3. However, within most applications r is typically small.

In another experiment, we set the values $r_1 = r_2 = 40$, $r = 60$, $m_1 = m_2 = 50000$, and vary the value of n from 500 to 3000 and compare the bounds given in (13) for finite completability with the one in (31) in Fig. 6. We note that the exploiting the multi-view structure leads to reduced sampling requirement.

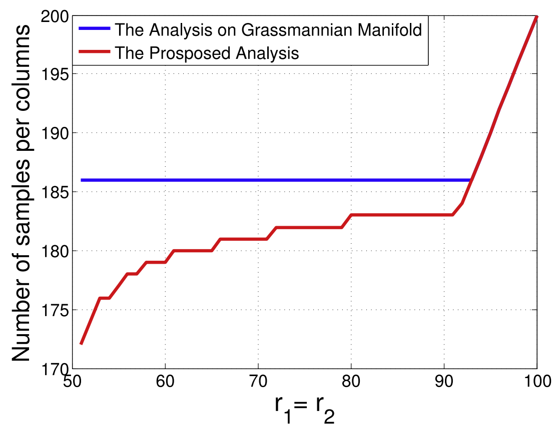
Note that we showed the difference of the bounds per column and this difference would be much more over all columns.



(a) $r = 40$.



(b) $r = 60$.



(c) $r = 100$.

Fig. 5. Lower bounds on the number of samples per column.

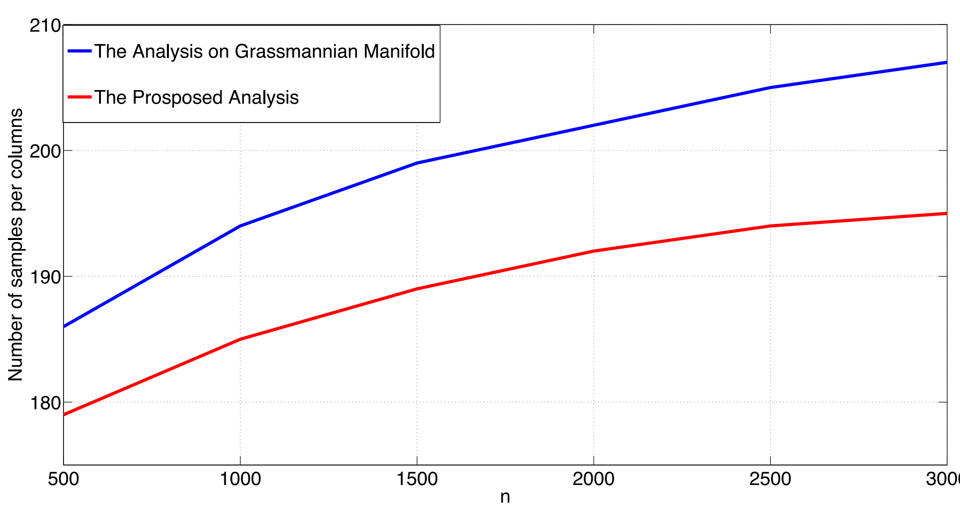


Fig. 6. Lower bounds on the number of samples per column.

7. Performance of completion algorithms

7.1. Multi-view matrix completion based on Newton's method

Note that the bounds on the sampling rate provided in previous sections (Lemmas 6 and 8) are the information theoretic bounds for unique/finite completability of the data. In other words, we found algorithm-independent guarantees for data completion. Since these fundamental bounds are obtained by analyzing the solvability of a set of polynomial equations obtained based on the sampled entries, it is then natural to expect that solving such a set of polynomial equations efficiently will lead to an efficient low-rank multi-view matrix completion algorithm, in the sense that completion is possible under very low-sampling rate (close to the information theoretic bound.) In fact, Newton's method has been employed to solve polynomial equation sets for matrix and tensor completion problems in [47]. Here we extend that method for multi-view matrix completion.

First, we show how to obtain the set of polynomials from the sampled entries using an example. Consider a matrix $\mathbf{U} \in \mathbb{R}^{3\times 6}$, where $\mathbf{U} = [\mathbf{U}_1|\mathbf{U}_2]$, $\mathbf{U}_1 \in \mathbb{R}^{3\times 3}$ (the first three columns) and $\mathbf{U}_2 \in \mathbb{R}^{3\times 3}$ (the last three columns). Assume that $r_1 = 2$, $r_2 = 2$ and $r = 3$. Moreover, suppose that the sampled entries of $\mathbf{U}$ are shown below.

$$\begin{bmatrix} \overbrace{- \quad 4.5 \quad -}^{\mathbf{U}_1} & \overbrace{- \quad - \quad 1.8}^{\mathbf{U}_2} \\ 1.4 \quad - \quad - & - \quad 10 \quad - \\ - \quad - \quad - & - \quad 1.2 \quad 5 \end{bmatrix}$$

Recall Eq. (1) and also $\mathbf{V} = [\mathbf{V}_1|\mathbf{V}_2|\mathbf{V}_3]$. For this example, since $r_1 = 2$, $r_2 = 2$ and $r = 3$, $\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3 \in \mathbb{R}^{3\times 1}$, and therefore they represent the first, second and the third columns of $\mathbf{V}$, respectively. Then, according to (2), we can obtain the following set of polynomials from the sampled entries

$$\mathbf{p}(\mathbf{z}) = \mathbf{0} \Rightarrow \begin{cases} V(1,1)T_1(1,2) + V(1,2)T_1(2,2) - 4.5 = 0, \\ V(2,1)T_1(1,1) + V(2,2)T_1(2,1) - 1.4 = 0, \\ V(1,2)T_2(1,3) + V(1,3)T_2(2,3) - 1.8 = 0, \\ V(2,2)T_2(1,2) + V(2,3)T_2(2,2) - 10 = 0, \\ V(3,2)T_2(1,2) + V(3,3)T_2(2,2) - 1.2 = 0, \\ V(3,2)T_2(1,3) + V(3,3)T_2(2,3) - 5 = 0, \end{cases}$$

where $\mathbf{z}$ denotes the vector of unknowns which in this case includes the entries of $\mathbf{V} \in \mathbb{R}^{nr}$, $\mathbf{T}_1 \in \mathbb{R}^{r_1 m_1}$ and $\mathbf{T}_2 \in \mathbb{R}^{r_2 m_2}$.

Let $\mathbf{z} \in \mathbb{R}^{(nr+r_1m_1+r_2m_2)\times 1}$ denote the vector that contains all the $(nr + r_1m_1 + r_2m_2)$ elements of the decomposition (1). Note that from (1) we know that each sampled entry of $\mathbf{U}_1$ ($\mathbf{U}_2$) results in a second-order polynomial that involves r_1 (r_2) entries of $\mathbf{V}$ and r_1 (r_2) entries of $\mathbf{T}_1$ ($\mathbf{T}_2$). As a result, we have a set of second-order polynomial equations $p_i(\mathbf{z}) = 0$, $i = 1, \dots, |\Omega|$, where $|\Omega|$ denotes the number of observed entries. Denote $\mathbf{p}(\mathbf{z}) = [p_1(\mathbf{z}), \dots, p_{|\Omega|}(\mathbf{z})]^\top$.

Remark 3. Note that finding $\mathbf{z}^* \in \mathbb{R}^{(nr+r_1m_1+r_2m_2)\times 1}$ such that $\mathbf{p}(\mathbf{z}^*) = \mathbf{0}$, is equivalent to finding a completion of the data.

In order to solve $\mathbf{p}(\mathbf{z}) = \mathbf{0}$, we use the well-known Newton's method. In particular, we start with some initial $\mathbf{z}_0 \in \mathbb{R}^{L\times 1}$, where $L = (nr + r_1m_1 + r_2m_2)$, and perform the following iteration

$$\mathbf{z}_n = \mathbf{z}_{n-1} - (\nabla \mathbf{p}(\mathbf{z}_{n-1}))^\dagger \mathbf{p}(\mathbf{z}_{n-1}), \tag{32}$$

where $\nabla \mathbf{p}(\mathbf{z}) \in \mathbb{R}^{|\Omega|\times L}$ and its (i, j)-th element denotes the partial derivative of $p_i(\mathbf{z})$ with respect to z_j, and the operator † denotes pseudoinverse.

We can easily observe that $\nabla \mathbf{p}(\mathbf{z})$ is a very sparse matrix. This is because the number of involved variables in each polynomial in $\mathbf{p(z)}$ is either $2r_1$ or $2r_2$. The sparse structure of this matrix enables a fast computation of its pseudoinverse, e.g., the command $\mathsf{sparse}(\nabla \mathbf{p}(\mathbf{z})) \setminus \mathbf{p}(\mathbf{z})$ in Matlab is an efficient way to calculate $(\nabla \mathbf{p}(\mathbf{z}_{n-1}))^\dagger \mathbf{p}(\mathbf{z}_{n-1})$.

7.2. Other completion methods for comparison

We will consider two alternative methods for completing the matrix, alternating minimization, and Newton's method for matrix completion. Both make only use of the rank value r, but not r_1 and r_2. In the alternating minimization approach, we can use the initialization scheme as in [15]. We compute the singular value decomposition (SVD) of $\mathbf{U}_{\boldsymbol{\Omega}}$, and choose the r largest eigenvalues and their corresponding eigenvectors as the initial $\mathbf{V}_0 \in \mathbb{R}^{n\times r}$ and $\mathbf{T}_0 \in \mathbb{R}^{r\times(m_1+m_2)}$. We use the rank decomposition $\mathbf{U} = \mathbf{VT}$ such that $\mathbf{V} \in \mathbb{R}^{n\times r}$ and $\mathbf{T} \in \mathbb{R}^{r\times(m_1+m_2)}$. Starting with the described initial $\mathbf{V}_0$ and $\mathbf{T}_0$, at the k-th iteration, given $\mathbf{V}_{k-1}$ and $\mathbf{T}_{k-1}$, we first update $\mathbf{V}_k$ by solving the following convex program

$$\text{minimize}_{\mathbf{V}_k \in \mathbb{R}^{n_1\times r}} \quad \|\mathbf{U}_\Omega - (\mathbf{V}_k\mathbf{T}_{k-1})_\Omega\|_{\mathcal{F}} \tag{33}$$

and then update $\mathbf{T}_k$ by solving

$$\text{minimize}_{\mathbf{T}_k \in \mathbb{R}^{r\times n_2}} \quad \|\mathbf{U}_\Omega - (\mathbf{V}_k\mathbf{T}_k)_\Omega\|_{\mathcal{F}} \tag{34}$$

where $\|\cdot\|_{\mathcal{F}}$ denotes the Frobenius norm.

Moreover, the Newton's method for matrix completion is described in [47], where we use only rank value r and derive the polynomial similar to multi-view case and use Newton's method to find a solution.

7.2.1. Initialization

For Newton's method, we simply use the initialization described in Section 7.2, by noting that $\mathbf{T}_1 = \mathbf{T}(1 : r_1, 1 : m_1)$ and $\mathbf{T}_2 = \mathbf{T}(r - r_2 + 1 : r, 1 : m_2)$. Then, by putting $\mathbf{V}_0$, $\mathbf{T}_0(1\!: r_1,\ 1\!: m_1)$ and $\mathbf{T}_0(r - r_2 + 1 : r, m_1 + 1 : m_1 + m_2)$ form the initialization in Section 7.2 and construct the initialization $\mathbf{z}_0$ for Newton's method.

7.2.2. Stopping criterion

In both of the described approaches, we stop the algorithm if either they converge or $\|\mathbf{z}_n\|$ (in Newton's method) and $\|\mathbf{V}_n\|$, $\|\mathbf{T}_n\|$ (in alternating minimization) become larger than $\max\{10^6, 10^6\|\mathbf{z}_0\|\}$ and $\max\{10^6, 10^6\|\mathbf{V}_0\|\}$, $\max\{10^6, 10^6\|\mathbf{T}_0\|\}$, respectively. In the case of divergence, we count this as a failure of the algorithm for recovering data.

7.3. Numericalexperiments for retrieving multi-view data

For the numerical experiments we consider an example where $n = 100, m_1 = 50, m_2 = 50, r_1 = r_2 = 5, r = 6$. Hence, $\mathbf{V}_1 \in \mathbb{R}^{100\times 1}$, $\mathbf{V}_2 \in \mathbb{R}^{100\times 4}$, and $\mathbf{V}_3 \in \mathbb{R}^{100\times 1}$. First, we generate random matrices $\mathbf{V} \in \mathbb{R}^{100\times 6}$, $\mathbf{T}_1 \in \mathbb{R}^{5\times 50}$ and $\mathbf{T}_2 \in \mathbb{R}^{5\times 50}$ (by choosing their entries independently and according to uniform distribution from [1,10]). Then, we construct $\mathbf{U}_1 = [\mathbf{V}_1|\mathbf{V}_2]\mathbf{T}_1 \in \mathbb{R}^{100\times 50}$ and $\mathbf{U}_2 = [\mathbf{V}_2|\mathbf{V}_3]\mathbf{T}_2 \in \mathbb{R}^{100\times 50}$. And finally construct $\mathbf{U} = [\mathbf{U}_1|\mathbf{U}_2] \in \mathbb{R}^{100\times 100}$. Then, we sample each entry independently with probability p and calculate the normalized number of samples by multiplying the number of samples to $L = (nr + r_1m_1 + r_2m_2)$ and dividing it by $n(m_1 + m_2)$. If $\frac{\|\hat{\mathbf{U}}-\mathbf{U}\|_{\mathcal{F}}}{\|\mathbf{U}\|_{\mathcal{F}}} < 0.01$, where $\hat{\mathbf{U}}$ denotes the obtained matrix through the corresponding completion algorithm, we count the experiment as a successful recovery. For each value of p, we complete 200 matrices and calculate the average recovery rate of each algorithm.

Note that the alternating minimization and Newton with matrix decomposition are only based on $n, (m_1 + m_2), r$ and do not take advantage of the other two rank values (r_1 and r_2). However, Newton with multi-view decomposition, which has the best performance makes use of multi-view decomposition and n, m_1, m_2, r, r_1, r_2 since the polynomials are derived from (1).

We have the following observations from Fig. 7:

1. The Newton's method for multi-view requires less number of samples for recovering the sampled data in comparison with the other methods. For example, when the normalized number of samples is equal to 2, it gives almost 100% recovery rate, while Newton's method for matrix completion gives 90% recovery and alternating minimization only recovers 20% of the scenarios. '
2. Alternating minimization is much worse than Newton's method in term of the required number of samples for recovery of the data.

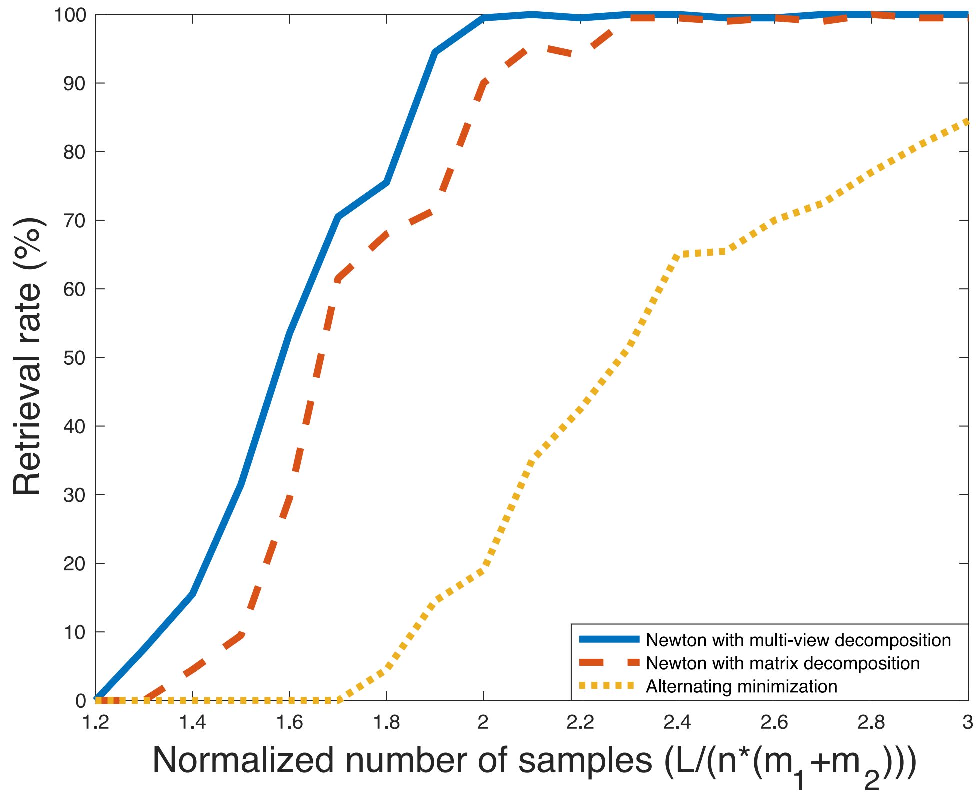


Fig. 7. Comparison of recovery rates.

3. Netwon's method for multi-view has small advantage in comparison with Newton's method for matrix, where one reason can be because we have used the same initialization for multi-view case.

8. Conclusions

This paper characterizes algorithm-independent conditions on the sampling pattern for finite completability of a low-rank multi-view matrix through an algebraic geometry analysis on the manifold structure of multi-view data. Then, having the mentioned analysis, we obtain the required number of sampled entries per column to guarantee that it leads to finite/unique completability with high probability. The numerical results demonstrate significant improvements in exploiting the multi-view data structure as compared to considering the two views separately. In other words, we obatined algorithm-independent guarantees for data completion. One important problem that we did not study in this paper and left it open for future is to develop an efficient algorithm that provably achieves the deterministic and probabilistic bounds derived in this paper. However, as an approach towards this goal, we developed the non-convex algorithm in Section 7 (similar to the approaches in [47] for matrix and tensor completion problems) for multi-view data by taking advantage of the rank decomposition and the manifold analysis in this work. Improving such non-convex optimization approaches to exactly achieve the mentioned bounds is left for future work.

Declaration of Competing Interest

The authors declare that they are only with Columbia University, Purdue University.

Acknowledgments

This work was supported in part by the U.S. National Science Foundation under Grant CCF-1814803 and in part by the U.S. Office of Naval Research under Grant N000141712827.

Appendix A. Proof of finite completability for the example in Section 2

Observe that Assumption 1 holds, i.e., each column of $\mathbf{U}_1$ includes at least one observed entry and each column of $\mathbf{U}_2$ includes at least two observed entries. According to the definition of the constraint matrix, we have $\breve{\mathbf{\Omega}} = [\breve{\mathbf{\Omega}}_1|\breve{\mathbf{\Omega}}_2]$, where

$$\breve{\mathbf{\Omega}}_1 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \text{and } \breve{\mathbf{\Omega}}_2 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Note that $r_1' = r - r_2 = 0$, $r_2' = r - r_1 = 1$ and $r' = r_1 + r_2 - r = 1$. As a result, $nr - r^2 - r_1^2 - r_2^2 + r(r_1 + r_2) = 5$ and $\breve{\mathbf{\Omega}}$ has exactly 5 columns. Suppose that $\mathbf{\Omega}'$ is an arbitrary submatrix of $\mathbf{\Omega}$. In order to show finite completability of $\mathbf{U}$, it suffices to show (6) holds. Let $\mathbf{\Omega}_1'$ and $\mathbf{\Omega}_2'$ denote the submatrix that consists of columns of $\breve{\mathbf{\Omega}}'$ that correspond to the first view and second view, respectively. Note that $\breve{\mathbf{\Omega}}' = [\breve{\mathbf{\Omega}}_1'|\breve{\mathbf{\Omega}}_2']$. Therefore, we only need to verify

$$(g(\breve{\mathbf{\Omega}}_2') - 2)^+ + (g(\breve{\mathbf{\Omega}}') - 1)^+ \geq c(\breve{\mathbf{\Omega}}'). \tag{35}$$

There are 3 different cases as follows:

1. $g(\breve{\mathbf{\Omega}}_2') = 0$: In this case, (35) reduces to $(g(\breve{\mathbf{\Omega}}_1') - 1)^+ \geq c(\breve{\mathbf{\Omega}}_1')$. This is easy to verify by checking each sub-case that $\breve{\mathbf{\Omega}}_1'$ has one or two columns of $\breve{\mathbf{\Omega}}_1$.
2. $g(\breve{\mathbf{\Omega}}_2') = 3$: In this case, (35) reduces to $1 + (g(\breve{\mathbf{\Omega}}') - 1)^+ \geq c(\breve{\mathbf{\Omega}}')$. We consider the following two sub-cases:

- $\breve{\boldsymbol{\Omega}}'_2$ is the first column of $\breve{\boldsymbol{\Omega}}_2$: Observe that in this case $c(\breve{\boldsymbol{\Omega}}') = c(\breve{\boldsymbol{\Omega}}'_1) + 1$, and also we always have $g(\breve{\boldsymbol{\Omega}}') \geq g(\breve{\boldsymbol{\Omega}}'_1)$. Hence, similar to the previous scenario, it suffices to show that $(g(\breve{\boldsymbol{\Omega}}'_1) - 1)^+ \geq c(\breve{\boldsymbol{\Omega}}'_1)$ which is easy to verify.
- $\breve{\boldsymbol{\Omega}}'_2$ does not include the first column of $\breve{\boldsymbol{\Omega}}_2$: Note that in this case $c(\breve{\boldsymbol{\Omega}}') \leq c(\breve{\boldsymbol{\Omega}}'_1) + 2$, and therefore it suffices to show that $(g(\breve{\boldsymbol{\Omega}}') - 1)^+ \geq c(\breve{\boldsymbol{\Omega}}'_1) + 1$. This is easy to verify by considering the fact that in this case $g(\breve{\boldsymbol{\Omega}}') = 4$ if and only if $\breve{\boldsymbol{\Omega}}'_1$ includes the second column of $\breve{\boldsymbol{\Omega}}_1$, and $g(\breve{\boldsymbol{\Omega}}') = 3$ otherwise.

3. $g(\breve{\boldsymbol{\Omega}}'_2) = 4$: In this case, (35) reduces to $2 + (g(\breve{\boldsymbol{\Omega}}') - 1)^+ \geq c(\breve{\boldsymbol{\Omega}}')$. Note that $g(\breve{\boldsymbol{\Omega}}'_2) = 4$ results that $g(\breve{\boldsymbol{\Omega}}') = 4$, and therefore (35) reduces to $5 \geq c(\breve{\boldsymbol{\Omega}}')$ which clearly always holds.

Appendix B. Example

We are interested to provide a non-trivial example to show that applying the existing matrix analysis for multi-view data (for each of the given rank constraints) can be very inefficient, while our analysis provides the necessary and sufficient condition for finite completability. This is just to show the efficiently of our analysis and also further clarifies the complicated statement and condition described in Theorem 1. In other words, this example shows how loose and inefficient the matrix analysis can be for multi-view data completion problem, while our proposed analysis is as efficient as possible by providing the necessary and sufficient condition.

Here we provide another motivating example such that $\max\{r_1, r_2\} < r < r_1 + r_2$. Consider a matrix $\mathbf{U} \in \mathbb{R}^{5\times 10}$, where $\mathbf{U} = [\mathbf{U}_1|\mathbf{U}_2]$, $\mathbf{U}_1 \in \mathbb{R}^{5\times 5}$ (the first five columns) and $\mathbf{U}_2 \in \mathbb{R}^{5\times 5}$ (the last five columns). Assume that $r_1 = 2$, $r_2 = 2$ and $r = 3$. Moreover, suppose that the sampled entries of $\mathbf{U}$ are shown with " $\times$ " below.

$$\mathbf{U} = \underbrace{\left[\begin{array}{ccccc|ccccc} \multicolumn{5}{c}{\overbrace{\hphantom{\times\ \times\ \times\ \times\ \times}}^{\mathbf{U}_1}} & \multicolumn{5}{c}{\overbrace{\hphantom{\times\ \times\ \times\ \times\ \times}}^{\mathbf{U}_2}} \\ \times & \times & \times & \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times & \times & \times & \times \\ \times & \times & - & - & - & \times & \times & - & - & - \\ - & - & \times & \times & - & - & - & \times & - & - \\ - & - & - & - & \times & - & - & - & \times & \times \end{array}\right]}_{\mathbf{U}}$$

We have the following observations about the number of completions of each matrix.

- *Given* $r_1 = 2$, $\mathbf{U}_1$ *is infinitely completable*: We can verify this via Lemma 1 since the fifth row of $\mathbf{U}_1$ has only one sampled entry.
- *Given* $r_2 = 2$, $\mathbf{U}_2$ *is infinitely completable*: We can verify this via Lemma 1 since the fourth row of $\mathbf{U}_2$ has only one sampled entry.
- *Given* $r = 3$, $\mathbf{U}$ *is infinitely completable*: We can easily verify this via Theorem 1 in [34].
- *For almost every matrix* $\mathbf{U}$, *given* $r_1 = 2$, $r_2 = 2$ *and* $r = 3$, $\mathbf{U}$ *is finitely completable*: Note that $nr - r^2 - r_1^2 - r_2^2 + r(r_1 + r_2) = 10$ and the constraint matrix includes exactly 10 columns. We can prove this statement by simply applying Theorem 1 which takes advantage of a geometric analysis on the manifold structure for multi-view data to incorporate all three rank constraints simultaneously.

Another approach to prove the finite completability given $r_1 = 2$, $r_2 = 2$ and $r = 3$ is as follows.

Note that the determinant of any 3×3 submatrix of $\mathbf{U}_1$ or $\mathbf{U}_2$ is 0 and due to the genericity assumption the determinant of any 2×2 submatrix of $\mathbf{U}_1$ or $\mathbf{U}_2$ is not 0 (with probability one). Hence, any 3×3 submatrix of $\mathbf{U}_1$ or $\mathbf{U}_2$ with 8 known entries results in obtaining the only unknown entry uniquely. Moreover, the determinant of any 4×4 submatrix of $\mathbf{U}$ is 0.

As a result, we can obtain the third and fourth rows of $\mathbf{U}_1$ uniquely. Moreover, we can obtain the third and fifth rows of $\mathbf{U}_2$ uniquely. Now, we show that the rest of the entries (the fifth row of $\mathbf{U}_1$ and the fourth row of $\mathbf{U}_2$) can be obtained uniquely as well. Define $\mathcal{I}_1 = \{1, 2, 3, 4\}$ and $\mathcal{I}_2 = \{1, 2, 7, 8\}$. Then, consider the 4×4 submatrix $\mathbf{U}(\mathcal{I}_1, \mathcal{I}_2)$ and note that it includes only one unknown entry, i.e., $\mathbf{U}(4, 7)$. Since the determinant of the 3×3 submatrix $\mathbf{U}(\mathcal{I}_1 \setminus \{4\}, \mathcal{I}_2 \setminus \{7\})$ is not 0 (with probability one), we can obtain $\mathbf{U}(4, 7)$ uniquely. It is easily verified that similarly the rest of the entries can be obtained uniquely.

Finally, we provide some observations for the described example to emphasize that removing any single sampled entry from this example leads to infinite completability with probability one. For example, removing $\mathbf{U}(4, 8)$ from the sampled entries results in infinite completability of the fourth row of $\mathbf{U}_2$. Define $\mathcal{I}_1 = \{1, 2, 3, 4\}$ and $\mathcal{I}'_2 = \{1, 2, 3, 6\}$. Note that even though the determinant of the 4×4 submatrix $\mathbf{U}(\mathcal{I}_1, \mathcal{I}'_2)$ is zero and the only unknown entry of this submatrix is $\mathbf{U}(4, 6)$, we cannot obtain $\mathbf{U}(4, 6)$ uniquely. This is because the determinant of the 3×3 submatrix $\mathbf{U}(\mathcal{I}_1 \setminus \{4\}, \mathcal{I}'_2 \setminus \{6\})$ is 0 (with probability one). Moreover, we can easily observe that removing $\mathbf{U}(3, 7)$ from the sampled entries results in infinite completability of the third and fourth rows of $\mathbf{U}_2$.

References

[1] D.L. Donoho, High-dimensional data analysis: the curses and blessings of dimensionality, in: AMS Math Challenges Lecture, 2000, pp. 1–32.
[2] C.R. Johnson, Matrix completion problems: a survey, in: Matrix theory and applications, 40, 1990, pp. 171–198. Providence, RI
[3] E.J. Candès, T. Tao, The power of convex relaxation: near-optimal matrix completion, IEEE Trans. Inform. Theory 56 (5) (2010) 2053.
[4] N. Subrahmanya, Y.C. Shin, Sparse multiple kernel learning for signal processing applications, IEEE Trans. Pattern Anal. Mach.Intell. 32 (5) (2010) 788–798.
[5] M. Liu, Y. Luo, D. Tao, C. Xu, Y. Wen, Low-rank multi-view learning in matrix completion for multi-label image classification, in: Proc. of Conference on Artificial Intelligence, 2015, pp. 2778–2784.
[6] W. Di, M.M. Crawford, View generation for multiview maximum disagreement based active learning for hyperspectral image classification, IEEE Trans. Geosci. Remote Sens. 50 (5) (2012) 1942–1954.
[7] L. Zhang, L. Zhang, D. Tao, X. Huang, On combining multiple features for hyperspectral remote sensing image classification, IEEE Trans. Geosci. Remote Sens. 50 (3) (2012) 879–893.
[8] M. Ashraphijuo, X. Wang, Clustering a union of low-rank subspaces of different dimensions with missing data, Pattern Recognit. Lett. 120 (2019) 31–35.
[9] D. Tao, X. Tang, X. Li, X. Wu, Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval, IEEE Trans. Pattern Anal. Mach.Intell. 28 (7) (2006) 1088–1099.
[10] A. Shon, K. Grochow, A. Hertzmann, R.P. Rao, Learning shared latent structure for image synthesis and robotic imitation, in: Proc. of Neural Information Processing Systems (NIPS), Advances in Neural Information Processing Systems, 2005, pp. 1233–1240.
[11] J. Yu, M. Wang, D. Tao, Semisupervised multiview distance metric learning for cartoon synthesis, IEEE Trans. Image Process. 21 (11) (2012) 4636–4648.
[12] S. Kiritchenko, S. Matwin, Email classification with co-training, in: Proc. of Conference of the Center for Advanced Studies on Collaborative Research (CASCON), IBM Corp., 2011, pp. 301–312.
[13] M. Ashraphijuo, X. Wang, V. Aggarwal, Rank determination for low-rank data completion, J.Mach. Learn. Res. 18 (1) (2017) 3422–3450.
[14] M. Amini, N. Usunier, C. Goutte, Learning from multiple partially observed views - an application to multilingual text categorization, in: Proc. of Neural Information Processing Systems (NIPS), 2009, pp. 28–36.
[15] P. Jain, P. Netrapalli, S. Sanghavi, Low-rank matrix completion using alternating minimization, in: Proc. of ACM Symposium on Theory of Computing (STOC), 2013, pp. 665–674.
[16] E.J. Candès, B. Recht, Exact matrix completion via convex optimization, Found. Comput. Math. 9 (6) (2009) 717–772.
[17] E.J. Candès, T. Tao, The power of convex relaxation: near-optimal matrix completion, IEEE Trans. Inf. Theory 56 (5) (2010) 2053–2080.
[18] M. Ashraphijuo, R. Madani, J. Lavaei, Characterization of rank-constrained feasibility problems via a finite number of convex programs, in: IEEE 55th Conference on Decision and Control (CDC), 2016, pp. 6544–6550.

[19] C. Xu, D. Tao, C. Xu, A survey on multi-view learning, arXiv:1304.5634 (2013).
[20] M. Gönen, E. Alpaydın, Multiple kernel learning algorithms, J. Mach. Learn. Res. 12 (2011) 2211–2268.
[21] F.R. Bach, G.R.G. Lanckriet, M.I. Jordan, Multiple kernel learning, conic duality, and the SMO algorithm, in: Proc. of International Conference on Machine Learning (ICML), 2004, p. 6.
[22] S. Bickel, T. Scheffer, Multi-view clustering, in: Proc. of IEEE International Conference on Data Mining (ICDM), 4, 2004, pp. 19–26.
[23] K. Chaudhuri, S.M. Kakade, K. Livescu, K. Sridharan, Multi-view clustering via canonical correlation analysis, in: Proc. of International Conference on Machine Learning (ICML), 2009, pp. 129–136.
[24] C. Christoudias, R. Urtasun, T. Darrell, Multi-view learning in the presence of view disagreement, arXiv:1206.3242 (2012).
[25] V. Sindhwani, D.S. Rosenberg, An RKHS for multi-view learning and manifold co-regularization, in: Proc. of International Conference on Machine Learning (ICML), 2008, pp. 976–983.
[26] P. Dhillon, D.P. Foster, L.H. Ungar, Multi-view learning of word embeddings via CCA, in: Proc. of Neural Information Processing Systems (NIPS), 2011, pp. 199–207.
[27] J. Farquhar, D. Hardoon, H. Meng, J.S. Shawe-taylor, S. Szedmak, Two view learning: SVM-2K, theory and practice, in: Proc. of Neural Information Processing Systems (NIPS), 2006, pp. 355–362.
[28] A. Kumar, P. Rai, H. Daume, Co-regularized multi-view spectral clustering, in: Proc. of Neural Information Processing Systems (NIPS), 2011, pp. 1413–1421.
[29] E.T. Matsubara, M.C. Monard, G.E. Batista, Multi-view semi-supervised learning: an approach to obtain different views from text datasets, in: Proc. of Conference on Advances in Logic Based Intelligent Systems (LAPTEC), 2005, pp. 97–104.
[30] S. Huang, Z. Kang, Z. Xu, Auto-weighted multi-view clustering via deep matrix decomposition, Pattern Recognit. 97 (2020).
[31] Y. Wang, L. Wu, X. Lin, J. Gao, Multiview spectral clustering via structured low-rank matrix factorization, IEEE Trans. Neural Netw. Learn.Syst. 29 (10) (2018) 4833–4843.
[32] W. Zhu, J. Lu, J. Zhou, Structured general and specific multi-view subspace clustering, Pattern Recognit. 93 (2019) 392–403.
[33] Q. Yin, J. Zhang, S. Wu, H. Li, Multi-view clustering via joint feature selection and partially constrained cluster label learning, Pattern Recognit. 93 (2019) 380–391.
[34] D. Pimentel-Alarcón, N. Boston, R. Nowak, A characterization of deterministic sampling patterns for low-rank matrix completion, IEEE J. Sel. Top. Signal Process. 10 (4) (2016) 623–636.
[35] F. Király, L. Theran, R. Tomioka, The algebraic combinatorial approach for low-rank matrix completion, J. Mach. Learn. Res. 16 (2015) 1391–1436.
[36] M. Ashraphijuo, X. Wang, Fundamental conditions on the sampling pattern for union of low-rank subspaces retrieval, Ann. Math. Artif. Intell. 87 (4) (2019) 373–393.
[37] M. Ashraphijuo, V. Aggarwal, X. Wang, Deterministic and probabilistic conditions for finite completability of low-tucker-rank tensor, IEEE Trans. Inf. Theory 65 (9) (2019) 5380–5400.
[38] M. Ashraphijuo, X. Wang, Fundamental conditions for low-CP-rank tensor completion, J. Mach. Learn. Res. 18 (63) (2017) 1–29.
[39] L. Balzano, B. Eriksson, R. Nowak, High rank matrix completion and subspace clustering with missing data, The Conference on Artificial Intelligence and Statistics (AIStats), 2012.
[40] D. Pimentel, R. Nowak, L. Balzano, On the sample complexity of subspace clustering with missing data, in: IEEE Workshop on Statistical Signal Processing (SSP), IEEE, 2014, pp. 280–283.
[41] D. Pimentel-Alarcón, E.R.D. Nowak, W. EDU, The information-theoretic requirements of subspace clustering with missing data, in: International Conference on Machine Learning, 2016.
[42] C. Yang, D. Robinson, R. Vidal, Sparse subspace clustering with missing entries, in: International Conference on Machine Learning (ICML), 2015, pp. 2463–2472.
[43] D. Pimentel-Alarcón, L. Balzano, R. Marcia, R. Nowak, R. Willett, Group-sparse subspace clustering with missing data, in: IEEE Workshop on Statistical Signal Processing (SSP), 2016, pp. 1–5.
[44] B. Sturmfels, Solving Systems of Polynomial Equations, no. 97, American Mathematical Society, 2002.
[45] M. Ashraphijuo, X. Wang, V. Aggarwal, A characterization of sampling patterns for low-rank multi-view data completion problem, in: 2017 IEEE International Symposium on Information Theory (ISIT), IEEE, 2017, pp. 1147–1151.
[46] D. Pimentel-Alarcón, N. Boston, R.D. Nowak, Deterministic conditions for subspace identifiability from incomplete sampling, in: IEEE International Symposium on Information Theory (ISIT), 2015, pp. 2191–2195.
[47] M. Ashraphijuo, X. Wang, J. Zhang, Low-rank data completion with very low sampling rate using Newton's method, IEEE Trans. Signal Process. 67 (7) (2019) 1849–1859.

Morteza Ashraphijuo received the B.Sc. degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2015, and the M.Sc. degree in industrial engineering and operations research from the University of California, Berkeley, CA, USA, in 2016. He is currently working toward the Ph.D. degree in electrical engineering at Columbia University, New York, NY, USA. His research interests include machine learning, optimization theory, and statistical learning. He was the recipient of multiple awards including a silver medal in the International Mathematical Olympiad (IMO) in 2010 and the Katta G. Murty Best Paper Prize from the University of California, Berkeley, in 2016.

Xiaodong Wang (S' 98-M' 98-SM' 04-F' 08) received the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, USA. He is currently a Professor of electrical engineering with Columbia University, New York, NY, USA. His research interests include computing, signal processing, and communications, and has authored and coauthored extensively in these areas. He has authored or coauthored a book entitled Wireless Communication Systems: Advanced Techniques for Signal Reception (Prentice Hall, 2003). His current research interests include wireless communications, statistical signal processing, and genomic signal processing. He was an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the IEEE TRANSACTIONS ON INFORMATION THEORY. He was the recipient of the 1999 NSF CAREER Award, the 2001 IEEE Communications Society and Information Theory Society Joint Paper Award, and the 2011 IEEE Communication Society Award for Outstanding Paper on New Communication Topics. He is listed as an ISI Highly Cited Author.

Vaneet Aggarwal (S'08-M'11-SM' 15) received the B.Tech. degree from the Indian Institute of Technology, Kanpur, India, and the M.A. and Ph.D. degrees from the Princeton University, Princeton, NJ, USA, all in electrical engineering, in 2005, 2007, and 2010, respectively. He is currently an Assistant Professor with the Purdue University, West Lafayette, IN, USA. Prior to this, he was a Senior Member of Technical Staff Research with AT&T Labs-Research, Middletown, NJ, USA, and an Adjunct Assistant Professor with the Columbia University, New York, NY, USA. His research interests include applications of information and coding theory to wireless systems and distributed storage systems. He was the recipient of the Princeton Universitys Porter Ogden Jacobus Honorific Fellowship in 2009.