

HIGH-DIMENSIONAL STATISTICAL INFERENCE: FROM
VECTOR TO MATRIX

Anru Zhang

A DISSERTATION

in

Applied Mathematics and Computational Science

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2015

Supervisor of Dissertation

T. Tony Cai
Dorothy Silberberg Professor of Statistics

Graduate Group Chairperson

Charles L. Epstein
Thomas A. Scott Professor of Mathematics

Dissertation Committee

T. Tony Cai, Professor
Dorothy Silberberg Professor of Statistics

Lawrance D. Brown
Miers Busch Professor of Statistics

Mark G. Low, Professor
Walter C. Bladstrom Professor of Statistics

Lawrance Shepp (in Memoriam)
Professor of Statistics

HIGH-DIMENSIONAL STATISTICAL INFERENCE: FROM VECTOR TO MATRIX

COPYRIGHT © 2015

Anru Zhang

Acknowledgments

I would like to first and foremost thank my advisor Professor Tony Cai. His sharp thinking and extensive knowledge have continuously been a great source of inspiration for my research. With his kindness, enthusiasm, patience, he guided me through a lot of hurdles during my PhD study. He is truly the epitome of an excellent scholar and mentor.

My sincere thanks go to Professor Larry Brown and Professor Mark Low for their kindly help and support for my research, teaching and job searching. Their brilliant comments and suggestions have been extremely helpful and will be valuable for my entire career. I want to thank the late Professor Larry Shepp. His smiles, jokes and encouragement will always be well remembered.

I am also very grateful to Professor Charles Epstein, who admitted me to this program and gave me such a great opportunity to study at UPenn. I want to thank my collaborators, Professor Tianxi Cai, Professor Dylan Small, Professor Hongzhe Li, Hyunseung Kang for the inspiring discussions I had with them. They have all given me generous support over the past five years.

I want to thank all the faculty members, staffs and my fellow students in both Department of Statistics and Program of Applied Mathematics and Computational

Science. Their accompany makes the past five years a very enjoyable and memorable experience.

My final, and most genuine acknowledgment goes to my family. Words cannot express how grateful I am to my parents and my wife Pixu, as they are my source of strength. They always gave me unconditional love and support whenever I met with difficulties. I would never be able to finish this thesis without them.

ABSTRACT

HIGH-DIMENSIONAL STATISTICAL INFERENCE: FROM VECTOR TO MATRIX

Anru Zhang

T. Tony Cai

Statistical inference for sparse signals or low-rank matrices in high-dimensional settings is of significant interest in a range of contemporary applications. It has attracted significant recent attention in many fields including statistics, applied mathematics and electrical engineering. In this thesis, we consider several problems including sparse signal recovery (compressed sensing under restricted isometry) and low-rank matrix recovery (matrix recovery via rank-one projections and structured matrix completion).

The first part of the thesis discusses compressed sensing and affine rank minimization in both noiseless and noisy cases and establishes sharp restricted isometry conditions for sparse signal and low-rank matrix recovery. The analysis relies on a key technical tool which represents points in a polytope by convex combinations of sparse vectors. The technique is elementary while leads to sharp results. It is shown that, in compressed sensing, $\delta_k^A < 1/3$, $\delta_k^A + \theta_{k,k}^A < 1$, or $\delta_{tk}^A < \sqrt{(t-1)/t}$ for any given constant $t \geq 4/3$ guarantee the exact recovery of all k sparse signals in the noiseless case through the constrained ℓ_1 minimization, and similarly in affine rank

minimization $\delta_r^M < 1/3$, $\delta_r^M + \theta_{r,r}^M < 1$, or $\delta_{tr}^M < \sqrt{(t-1)/t}$ ensure the exact reconstruction of all matrices with rank at most r in the noiseless case via the constrained nuclear norm minimization. Moreover, for any $\epsilon > 0$, $\delta_k^A < 1/3 + \epsilon$, $\delta_k^A + \theta_{k,k}^A < 1 + \epsilon$, or $\delta_{ik}^A < \sqrt{\frac{t-1}{t}} + \epsilon$ are not sufficient to guarantee the exact recovery of all k -sparse signals for large k . Similar result also holds for matrix recovery. In addition, the conditions $\delta_k^A < 1/3$, $\delta_k^A + \theta_{k,k}^A < 1$, $\delta_{ik}^A < \sqrt{(t-1)/t}$ and $\delta_r^M < 1/3$, $\delta_r^M + \theta_{r,r}^M < 1$, $\delta_{tr}^M < \sqrt{(t-1)/t}$ are also shown to be sufficient respectively for stable recovery of approximately sparse signals and low-rank matrices in the noisy case.

For the second part of the thesis, we introduce a rank-one projection model for low-rank matrix recovery and propose a constrained nuclear norm minimization method for stable recovery of low-rank matrices in the noisy case. The procedure is adaptive to the rank and robust against small perturbations. Both upper and lower bounds for the estimation accuracy under the Frobenius norm loss are obtained. The proposed estimator is shown to be rate-optimal under certain conditions. The estimator is easy to implement via convex programming and performs well numerically. The techniques and main results developed in the chapter also have implications to other related statistical problems. An application to estimation of spiked covariance matrices from one-dimensional random projections is considered. The results demonstrate that it is still possible to accurately estimate the covariance matrix of a high-dimensional distribution based only on one-dimensional projections.

For the third part of the thesis, we consider another setting of low-rank matrix completion. Current literature on matrix completion focuses primarily on independent sampling models under which the individual observed entries are sampled independently. Motivated by applications in genomic data integration, we propose a new framework of structured matrix completion (SMC) to treat structured missingness by design. Specifically, our proposed method aims at efficient matrix recovery when

a subset of the rows and columns of an approximately low-rank matrix are observed. We provide theoretical justification for the proposed SMC method and derive lower bound for the estimation errors, which together establish the optimal rate of recovery over certain classes of approximately low-rank matrices. Simulation studies show that the method performs well in finite sample under a variety of configurations. The method is applied to integrate several ovarian cancer genomic studies with different extent of genomic measurements, which enables us to construct more accurate prediction rules for ovarian cancer survival.

Contents

1	Compressed Sensing under Restricted Isometry	5
1.1	Introduction	5
1.2	Compressed Sensing	13
1.3	Affine Rank Minimization	20
1.4	Discussions	24
2	ROP: Matrix Recovery via Rank-one Projections	28
2.1	Introduction	28
2.2	Matrix Recovery under Gaussian Noise	34
2.3	Sub-Gaussian Design and Sub-Gaussian Noise	46
2.4	Application to Estimation of Spiked Covariance Matrix	48
2.5	Simulation Results	51
2.6	Discussions	59
3	Structured Matrix Completion	62
3.1	Introduction	62
3.2	Structured Matrix Completion: Methodology	67
3.3	Theoretical Analysis	76

3.4	Simulation	82
3.5	Application in Genomic Data Integration	88
3.6	Discussions	93
4	Future Problems	95
4.1	Inference for Large Gaussian Graphical Model with Missing Data	95
4.2	High-dimensional Sparsity Test	96
4.3	Noisy Structured Matrix Completion	96
A	Appendices	97
A.1	Supplement for Chapter 1	98
A.2	Supplement for Chapter 2	145
A.3	Supplement for Chapter 3	191
	Bibliography	226

List of Tables

3.1	Relative spectral norm loss ($\ \hat{A}_{22} - A_{22}\ /\ A_{22}\ $) and Frobenius norm loss ($\ \hat{A}_{22} - A_{22}\ _F/\ A_{22}\ _F$) for $p_1 = 1148$, $p_2 = 1225$, $m_1 = 230$, $m_2 = 426$ and singular values of A being $\{j^{-\alpha} : j = 1, 2, \dots\}$	88
3.2	Shown in (a) are the estimates of the log hazard ratio (logHR) along with their corresponding standard errors (SE) and p -values by fitting stratified Cox model integrating information from 4 independent studies with imputed miRNA based on the SMC method and the nuclear norm minimization (NNM); and Cox model to the TCGA test data with original observed miRNA (Ori). Shown also are the estimates for each individual studies by fitting separate Cox models with imputed miRNA.	92

List of Figures

1.1	A graphical illustration of sparse representation of a polytope in one orthant with $p = 3$ and $s = 2$. All the points in the colored area can be expressed as convex combinations of the sparse vectors represented by the three pointed black line segments on the edges.	11
1.2	Plot of δ_* as a function of t . The dotted line is $t = 4/3$	25
1.3	Plot of n^* as a function of t	27
2.1	Rates of successful recovery for the ROP and Gaussian ensemble with $p_1 = p_2 = 100$, $r = 5$, and $n = Cr \max(p_1, p_2)$ for C ranging from 3 to 6.	52
2.2	Recovery accuracy (solid line) for approximately low-rank matrices with different values of r , where $p_1 = p_2 = 100$, $n = 2000$, $\sigma(A) = (1, 1/\sqrt{2}, \dots, 1/\sqrt{r})$. The dashed line is the theoretical upper bound.	54
2.3	Left Panel: Comparison of the proposed estimator with \hat{A}^{ℓ_1} and \hat{A}^{DS} for $p_1 = p_2 = 50$, $r = 4$, $\sigma = 0.01$, and n ranging from 850 to 1200. Right Panel: Ratio of the squared Frobenius norm loss of \hat{A}^{ℓ_1} to that of the proposed estimator for $p_1 = p_2 = 50$, $r = 4$, and n varying from 2000 to 15000.	55

2.4	Comparison of the proposed estimator \hat{A} with the \hat{A}^{ℓ_1} . Here $p = 40$, $r = 5$, $\sigma = 0.1, 0.01, 0.001, 0.0001$, and n ranges from 50 to 800.	56
2.5	Comparison of the performance with cross validation and without cross-validation in both ROP and SROP. Left panel: ROP, $p_1 = p_2 = 30$, $r = 4$, n varies from 750 to 1400. Right panel: SROP, $p = 40$, $r = 5$, n varies from 50 to 800.	58
2.6	Original grayscale MIT logo	59
2.7	recovery of MIT logo based on different number of measurements. Left: 900; Middle: 1000; Right: 1080.	60
3.1	Illustrative example with $A \in \mathbb{R}^{30 \times 30}$, $m_1 = m_2 = 10$. (A darker block corresponds to larger magnitude.)	72
3.2	Searching for the appropriate position to truncate from $\hat{r} = 10$ to 1.	73
3.3	Spectral norm loss (left panel) and Frobenius norm loss (right panel) when there is a gap between $\sigma_r(A)$ and $\sigma_{r+1}(A)$. The singular value values of A are given by (3.21), $p_1 = p_2 = 1000$, and $m_1 = m_2 = 50$	84
3.4	Spectral norm loss (left panel) and Frobenius norm loss (right panel) as the thresholding constant c varies. The singular values of A are $\{j^{-\alpha}, j = 1, 2, \dots\}$ with α varying from 0.3 to 2, $p_1 = p_2 = 1000$, and $m_1 = m_2 = 50$	85
3.5	Spectral and Frobenius norm losses with column/row thresholding. The singular values of A are $\{j^{-1}, j = 1, 2, \dots\}$, $p_1 = 300$, $p_2 = 3000$, and $m_1, m_2 = 10, \dots, 150$	86
3.6	Comparison of the proposed SMC method with the NNM method with 5-cross-validation for the settings with singular values of A being $\{j^{-\alpha}, j = 1, 2, \dots\}$ for α ranging from 0.6 to 2, $p_1 = p_2 = 500$, and $m_1 = m_2 = 50$ or 100.	87

3.7	Imputation scheme for integrating multiple OC genomic studies. . . .	90
A.1	Losses for the settings with singular values of A being $\{j^{-1}, j = 1, 2, \dots\}$, $p_1 = p_2 = 1000, m_1, m_2 = 10, \dots, 210$	191
A.2	Losses for settings with singular values of A being $\{j^{-1}, j = 1, 2, 3, \dots\}$, $p_2 = 1000, m_2 = 50, m_1/p_1 = 1/4, 1/12, 1/20, 1/28, 1/36$, and $p_1 =$ $100, \dots, 100,000$	192

Preface

High-dimensional statistical inference has been a very active area in the recent years. During my graduate studies, I have been fascinated by a range of interesting problems in this field. These problems are motivated by important applications in many areas, from genomics to signal processing to social networks to climate studies, and by considerations from statistical theory. Some of these problems exhibit new features that are very different from those in the conventional low-dimensional settings. In this thesis, we discuss some recent advances on several problems in high-dimensional statistical inference, including sparse signal recovery (compressed sensing) and low-rank matrix recovery. Before we elaborate these problems respectively in Chapter 1 – 3, the quick overviews are provided below.

Compressed Sensing under Restricted Isometry¹

Efficient recovery of sparse signals, or compressed sensing, has been a very active area of recent research in applied mathematics, statistics, and machine learning, with many important applications, ranging from signal processing to medical imaging to radar systems. A central goal is to develop fast algorithms that can recover sparse

¹This part of the thesis is published in Cai and Zhang (2013a,b, 2014b).

signals from a relatively small number of linear measurements.

Among different frameworks and methods for compressed sensing, the restricted isometry property (RIP) and the constrained ℓ_1 norm minimization are very well-known and widely used. To be specific, under the RIP framework, we use the restricted isometry constant, δ_k , and restricted orthogonal constant, θ_{k_1, k_2} , to regularize the sensing matrix. Previous literature has shown that as long as δ_k and θ_{k_1, k_2} are small, the exact recovery of sparse signal in the noiseless case and stable recovery in the noisy case can be guaranteed by the ℓ_1 minimization. However, a fundamental question for the RIP framework is that of how small it is necessary for δ_k or θ_{k_1, k_2} to be.

In Chapter 1, we show that if any one of the following conditions is met, (1) $\delta_k < 1/3$, (2) $\delta_k + \theta_{k,k} < 1$, (3) $\delta_{tk} < \sqrt{(t-1)/t}$ for some $t > 4/3$, the exact recovery and stable recovery of all k -sparse signals can be guaranteed by using ℓ_1 minimization in the noiseless case and noisy case, respectively. On the other hand, we further prove that the bounds $\delta_k < 1/3$, $\delta_k + \theta_{k,k} < 1$ and $\delta_{tk} < \sqrt{(t-1)/t}$ are sharp in the sense that there exist sensing matrices such that it is impossible to recover all k -sparse signals accurately but either (1) $\delta_k < 1/3 + \varepsilon$, or (2) $\delta_k + \theta_{k,k} < 1 + \varepsilon$ or (3) $\delta_{tk} < \sqrt{(t-1)/t} + \varepsilon$ holds for a small value $\varepsilon > 0$. It is also shown that the same results hold for low-rank matrix recovery under the trace regression model.

Meanwhile, we also develop a useful technical geometric tool which represents points in a high-dimensional polytope by convex combinations of sparse vectors (Lemma 1.1.1) and this is of independent interest.

Matrix Recovery via Rank-One Projections²

We introduce a rank-one projection (ROP) model for low-rank matrix recovery and propose a new convex constrained minimization method for stable recovery of low-rank matrices in the noisy case. The procedure is adaptive to the rank and robust against small perturbations. Both upper and lower bounds for the estimation accuracy under the Frobenius norm loss are obtained. The proposed estimator is shown to be rate-optimal under certain conditions. The estimator is easy to implement via convex programming and performs well numerically. Compared to some of the other frameworks in the literature (e.g. Gaussian ensemble or matrix completion), the proposed procedure requires only a small amount of storage space and can recover all low-rank matrices with no additional structural assumptions.

The techniques and main results developed for ROP also have implications to other related statistical problems. An application to estimation of spiked covariance matrices from one-dimensional random projections is also considered. The results demonstrate, somewhat surprisingly, that it is still possible to accurately estimate the covariance matrix of a high-dimensional distribution based only on one-dimensional projections.

Structured Matrix Completion³

In some other applications such as genomic data integration and paleoclimate reconstruction, the model is highly structured in a way that observed entries are all either in full rows or full columns. In other words, the rows and columns can be permuted so that the missing part of the matrix becomes a contiguous block. In this case, some well-studied matrix recovery methods, such as penalized nuclear norm minimization

²This part of thesis is published in Cai and Zhang (2015).

³This part of thesis is published in Cai et al. (2015).

or constraint nuclear norm minimization, are shown to be inappropriate.

In Chapter 3, we propose a new framework of structured matrix completion (SMC) to treat this structured missingness by design. The new SMC method, whose main idea is based on the Schur Complement, can be easily implemented by a fast algorithm which only involves basic matrix operations and the singular value decomposition. We also provide theoretical justification for the proposed SMC method and derive lower bounds for the estimation errors. These together establish the optimal rate of recovery over certain classes of approximately low-rank matrices. Both theoretical and numerical studies show that SMC recovers low-rank matrices accurately and is robust against small perturbations. This method was also applied to integrate several ovarian cancer genomic studies with different extent of genomic measurements, which enables us to construct more accurate prediction rules for ovarian cancer survival.

Compressed Sensing under Restricted Isometry

1.1 Introduction

Efficient recovery of sparse signals and low-rank matrices has been a very active area of recent research in applied mathematics, statistics, and machine learning, with many important applications, ranging from signal processing (Tropp et al. 2010, Davenport et al. 2012) to medical imaging (Lustig et al., 2008) to radar systems (Baraniuk and Steeghs 2007, Herman and Strohmer 2009). A central goal is to develop fast algorithms that can recover sparse signals and low-rank matrices from a relatively small number of linear measurements. Constrained ℓ_1 -norm minimization and nuclear norm minimization are among the most well-known algorithms for the recovery of sparse signals and low-rank matrices respectively.

In compressed sensing, one observes

$$y = A\beta + z, \tag{1.1}$$

where $y \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times p}$ with $n \ll p$, $\beta \in \mathbb{R}^p$ is an unknown sparse signal, and $z \in \mathbb{R}^n$ is a vector of measurement errors. The goal is to recover the unknown signal $\beta \in \mathbb{R}^p$ based on the measurement matrix A and the observed signal y .

For the reconstruction of β , the most intuitive approach is to find the sparsest signal in the feasible set of possible solutions, i.e.,

$$\text{minimize } \|\beta\|_0, \text{ subject to } A\beta - y \in \mathcal{B}$$

where $\|\beta\|_0$ denote the ℓ_0 norm of β , which is defined to be the number of nonzero coordinates, and \mathcal{B} is a bounded set determined by the error structure. However, it is well-known that this method is NP-hard and thus computationally infeasible in the high dimensional settings. Convex relaxations of this method has been proposed and studied in the literature. The constrained ℓ_1 minimization method proposed by Candès and Tao (2005) estimates the signal β by

$$\hat{\beta} = \arg \min_{\beta \in \mathbb{R}^p} \{\|\beta\|_1 : \text{subject to } A\beta - y \in \mathcal{B}\}, \quad (1.2)$$

where \mathcal{B} is a set determined by the noise structure. In particular, \mathcal{B} is taken to be $\{0\}$ in the noiseless case. This constrained ℓ_1 minimization method has now been well studied and it is understood that the procedure provides an efficient method for sparse signal recovery.

A closely related problem to compressed sensing is the affine rank minimization problem (ARMP) (Recht et al., 2010), which aims to recover an unknown low-rank matrix based on its affine transformation. In ARMP, one observes

$$b = \mathcal{M}(X) + z, \quad (1.3)$$

where $\mathcal{M} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^q$ is a known linear map, $X \in \mathbb{R}^{m \times n}$ is an unknown low-rank matrix of interest, and $z \in \mathbb{R}^q$ is measurement error. The goal is to recover the low-rank matrix X based on the linear map \mathcal{M} and the observation $b \in \mathbb{R}^q$.

To recover X , the most intuitive approach is to find the the lowest-rank matrix in the feasible set of possible solutions, i.e.,

$$\text{minimize } \text{rank}(X), \text{ subject to } \mathcal{M}(X) - y \in \mathcal{B},$$

Similarly to the ℓ_0 norm minimization in compressed sensing, the rank minimization is also NP-hard and thus computationally infeasible in the high dimensional settings. Constrained nuclear norm minimization (Recht et al., 2010), which is analogous to ℓ_1 minimization in compressed sensing, estimates X by

$$X_* = \arg \min_{B \in \mathbb{R}^{m \times n}} \{\|B\|_* : \text{subject to } \mathcal{M}(B) - b \in \mathcal{B}\}, \quad (1.4)$$

where $\|B\|_*$ is the nuclear norm of B , which is defined as the sum of all singular values of B .

One of the most widely used frameworks in compressed sensing is the *restrict isometry property* (RIP) introduced in Candès and Tao (2005). A vector $\beta \in \mathbb{R}^p$ is called s -sparse if $|\text{supp}(\beta)| \leq s$, where $\text{supp}(\beta) = \{i : \beta_i \neq 0\}$ is the support of β .

Definition 1.1.1. *Let $A \in \mathbb{R}^{n \times p}$ and let $1 \leq k, k_1, k_2 \leq p$ be integers. The restricted isometry constant (RIC) of order k is defined to be the smallest non-negative number δ_k^A such that*

$$(1 - \delta_k^A)\|\beta\|_2^2 \leq \|A\beta\|_2^2 \leq (1 + \delta_k^A)\|\beta\|_2^2 \quad (1.5)$$

for all k -sparse vectors β . The restricted orthogonality constant (ROC) of order (k_1, k_2) is defined to be the smallest non-negative number θ_{k_1, k_2}^A such that

$$|\langle A\beta_1, A\beta_2 \rangle| \leq \theta_{k_1, k_2}^A \|\beta_1\|_2 \|\beta_2\|_2 \quad (1.6)$$

for all k_1 -sparse vector β_1 and k_2 -sparse vector β_2 with disjoint supports.

Similar to the RIP for the measurement matrix A in compressed sensing given in Definition 1.1.1, a restricted isometry property for a linear map \mathcal{M} in ARMP can be given. For two matrices X and Y in $\mathbb{R}^{m \times n}$, define their inner product as $\langle X, Y \rangle = \sum_{i,j} X_{ij} Y_{ij}$ and the Frobenius norm as $\|X\|_F = \sqrt{\langle X, X \rangle} = \sqrt{\sum_{i,j} X_{ij}^2}$.

Definition 1.1.2. Let $\mathcal{M} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^p$ be a linear map and let $1 \leq r, r_1, r_2 \leq \min(m, n)$ be integers. The restricted isometry constant (RIC) of order r is defined to be the smallest non-negative number $\delta_r^{\mathcal{M}}$ such that

$$(1 - \delta_r^{\mathcal{M}}) \|X\|_F^2 \leq \|\mathcal{M}(X)\|_2^2 \leq (1 + \delta_r^{\mathcal{M}}) \|X\|_F^2 \quad (1.7)$$

for all $m \times n$ matrix X of rank at most r . The restricted orthogonality constant (ROC) of order (r_1, r_2) is defined to be the smallest non-negative number $\theta_{r_1, r_2}^{\mathcal{M}}$ such that

$$|\langle \mathcal{M}(X_1), \mathcal{M}(X_2) \rangle| \leq \theta_{k_1, k_2}^{\mathcal{M}} \|X_1\|_F \|X_2\|_F \quad (1.8)$$

for all matrices X_1 and X_2 which have rank at most r_1 and r_2 respectively, and satisfy $X_1^T X_2 = 0$ and $X_1 X_2^T = 0$.

In addition to RIP, another widely used criterion is the mutual incoherence property (MIP) defined in terms of $\mu = \max_{i \neq j} |\langle A_i, A_j \rangle|$. See, for example, Donoho and Huo (2001) and Cai et al. (2010d). The MIP is a special case of the restricted orthogonal property as $\mu = \theta_{1,1}$ when the columns of A are normalized.

Roughly speaking, the RIC δ_k^A and ROC θ_{k_1, k_2}^A measure how far subsets of cardinality k and k_1, k_2 of columns of A are to an orthonormal system. It is obvious that δ_k and θ_{k_1, k_2} are increasing in each of their indices. It is noteworthy that our definition of ROC in the matrix case is different from the one given in Mohan and Fazel (2010).

Different conditions on the RIC and ROC for sparse signal recovery have been introduced and studied in the literature. For example, sufficient conditions for the

exact recovery in the noiseless case include $\delta_{2k} < \sqrt{2} - 1$ in Candès (2008), $\delta_{2k} < 0.472$ in Cai et al. (2010c), $\delta_{2k} < 0.497$ in Mo and Li (2011), $\delta_k < 0.307$ in Cai et al. (2010b), $\delta_{2k}^A < 4/\sqrt{41}$ in Andersson and Stromberg (2014), $\delta_k^A + \theta_{k,k}^A + \theta_{k,2k}^A < 1$ Candès and Tao (2005); $\delta_{2k}^A + \theta_{k,2k}^A < 1$ Candès and Tao (2007); $\delta_{1.5k}^A + \theta_{k,1.5k}^A < 1$ Cai et al. (2009), $\delta_{1.25k}^A + \theta_{k,1.25k}^A < 1$ Cai et al. (2010c), and $\theta_{1,1}^A < \frac{1}{2k-1}$ when $\delta_1^A = 0$ (Donoho and Huo 2001, Fuchs 2004, Cai et al. 2010d). There are also other sufficient conditions that involve RICs of different orders, e.g. $\delta_{3k}^A + 3\delta_{4k}^A < 2$ in Candès et al. (2006), $\delta_{2k}^A < 0.5746$ jointly with $\delta_{8k}^A < 1$, $\delta_{3k}^A < 0.7731$ jointly with $\delta_{16k}^A < 1$ in Zhou et al. (2013). As in compressed sensing, there are many sufficient conditions based on the RIC to guarantee the exact recovery of matrices of rank at most r through the constrained nuclear norm minimization (1.4). These include $\delta_{4r}^M < \sqrt{2} - 1$ in Candès and Plan (2011), $\delta_{5r}^M < 0.607$, $\delta_{4r}^M < 0.558$, and $\delta_{3r}^M < 0.4721$ in Mohan and Fazel (2010), $\delta_{2r}^M < 0.4931$ and $\delta_r^M < 0.307$ in Wang and Li (2013), $\delta_{2r+\alpha r} + \frac{1}{\sqrt{\beta}}\theta_{2r+\alpha r, \beta r} < 1$ where $2\alpha \leq \beta \leq 4\alpha$ in Mohan and Fazel (2010). It is however unclear if any of these conditions can be further improved.

In this chapter, we develop a new elementary technique for the analysis of the constrained ℓ_1 -norm minimization and nuclear norm minimization procedures and establish sharp RIP conditions on RICs and ROCs for sparse signal and low-rank matrix recovery. The analysis is surprisingly simple, while leads to sharp results. The key technical tool we develop states an elementary geometric fact: Any point in a polytope can be represented as a convex combination of sparse vectors. The following lemma may be of independent interest.

Lemma 1.1.1 (Sparse Representation of a Polytope). *For a positive number α and a positive integer s , define the polytope $T(\alpha, s) \subset \mathbb{R}^p$ by*

$$T(\alpha, s) = \{v \in \mathbb{R}^p : \|v\|_\infty \leq \alpha, \|v\|_1 \leq s\alpha\}.$$

For any $v \in \mathbb{R}^p$, define the set of sparse vectors $U(\alpha, s, v) \subset \mathbb{R}^p$ by

$$\begin{aligned} U(\alpha, s, v) = \{ & u \in \mathbb{R}^p : \text{supp}(u) \subseteq \text{supp}(v), \|u\|_0 \leq s, \\ & \|u\|_1 = \|v\|_1, \|u\|_\infty \leq \alpha \}. \end{aligned} \quad (1.9)$$

Then $v \in T(\alpha, s)$ if and only if v is in the convex hull of $U(\alpha, s, v)$. In particular, any $v \in T(\alpha, s)$ can be expressed as

$$v = \sum_{i=1}^N \lambda_i u_i, \quad \text{and } 0 \leq \lambda_i \leq 1, \quad \sum_{i=1}^N \lambda_i = 1, \quad \text{and } u_i \in U(\alpha, s, v).$$

Lemma 1.1.1 shows that any point $v \in \mathbb{R}^p$ with $\|v\|_\infty \leq \alpha$ and $\|v\|_1 \leq s\alpha$ must lie in a convex polytope whose extremal points are s -sparse vectors u with $\|u\|_1 = \|v\|_1$ and $\|u\|_\infty \leq \alpha$, and vice versa. This geometric fact turns out to be a powerful tool in analyzing constrained ℓ_1 -norm minimization for compressed sensing and nuclear norm minimization for ARMP, since it represents a non-sparse vector by the sparse ones, which provides a bridge between general vectors and the RIP conditions. A graphical illustration of Lemma 1.1.1 is given in Figure 1.1.

Combining the results developed in Sections 1.2 and 1.3, we establish the following sharp sufficient RIP conditions for the exact recovery of all k -sparse signals and low-rank matrices in the noiseless case. We focus here on the exact sparse (low-rank) and noiseless case; the general approximately sparse (low-rank) and noisy case is considered in Sections 1.2 and 1.3.

Theorem 1.1.1. *Let $y = A\beta$ where $\beta \in \mathbb{R}^p$ is a k -sparse vector. If any of the following conditions hold*

1. $\delta_k^A < 1/3$,
2. $\delta_k^A + \theta_{k,k}^A < 1$,

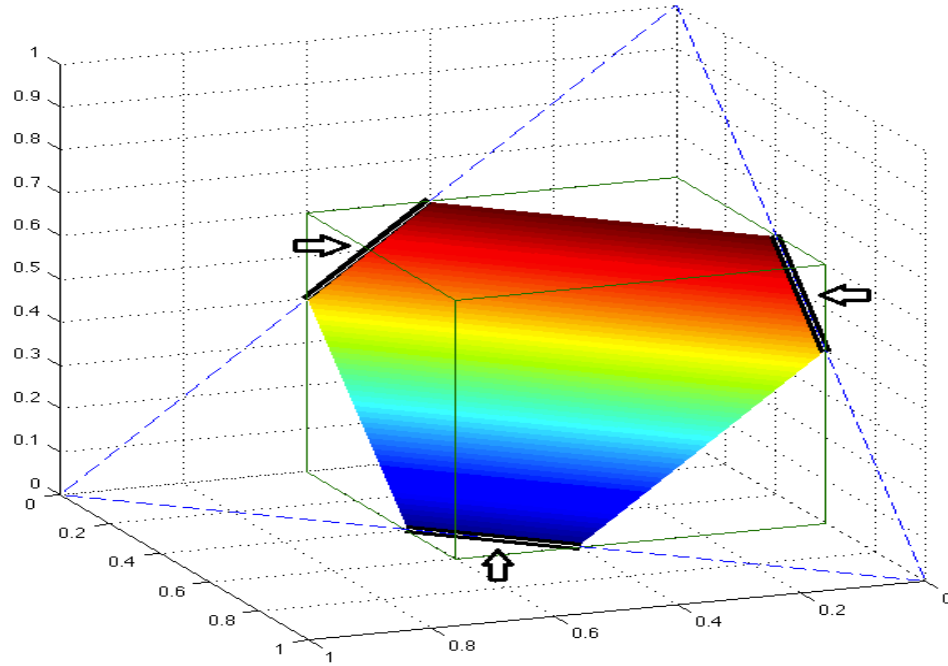


Figure 1.1: A graphical illustration of sparse representation of a polytope in one orthant with $p = 3$ and $s = 2$. All the points in the colored area can be expressed as convex combinations of the sparse vectors represented by the three pointed black line segments on the edges.

$$3. \delta_{tk}^A < \sqrt{\frac{t-1}{t}} \text{ for some } t \geq 4/3,$$

then the ℓ_1 norm minimizer $\hat{\beta}$ of (1.2) with $\mathcal{B} = \{0\}$ recovers β exactly.

Similarly, suppose $b = \mathcal{M}(X)$ where the matrix $X \in \mathbb{R}^{m \times n}$ is of rank at most r .

If any of the following conditions hold

$$1. \delta_r^{\mathcal{M}} < 1/3,$$

$$2. \delta_r^{\mathcal{M}} + \theta_{r,r}^{\mathcal{M}} < 1,$$

$$3. \delta_{tr}^{\mathcal{M}} < \sqrt{\frac{t-1}{t}} \text{ for some } t \geq 4/3,$$

then the nuclear norm minimizer X_* of (1.4) with $\mathcal{B} = \{0\}$ recovers X exactly.

Moreover, it will be shown that for any $\epsilon > 0$, $\delta_k^A < \frac{1}{3} + \epsilon$, $\delta_k^A + \theta_{k,k}^A < 1 + \epsilon$, or $\delta_{tk}^A < \sqrt{\frac{t-1}{t}} + \epsilon$ are not sufficient to guarantee the exact recovery of all k -sparse signals for large k . Similar results also hold for matrix recovery. For the more general approximately sparse (low-rank) and noisy cases considered in Sections 1.2 and 1.3, it is shown that the conditions in Theorem 1.1.1 are also sufficient respectively for stable recovery of (approximately) k -sparse signals and (approximately) rank- r matrices in the noisy case. An oracle inequality is also given in the case of compressed sensing with Gaussian noise under the conditions $\delta_k^A < 1/3$, $\delta_k^A + \theta_{k,k}^A < 1$ and $\delta_{tk}^A < \sqrt{(t-1)/t}$ when $t \geq 4/3$.

The rest of the chapter is organized as follows. Section 1.2 considers sparse signal recovery and Section 1.3 focuses on low-rank matrix recovery. Discussions on the case $t < 4/3$ and some related issues are given in Section 1.4. The proofs of the key technical result Lemma 1.1.1 and the main theorems are contained in the Appendix A.1.

1.2 Compressed Sensing

We consider compressed sensing in this section and establish the sufficient RIP condition $\delta_k^A < 1/3$, $\delta_k^A + \theta_{k,k}^A < 1$ and $\delta_{tk}^A < \sqrt{(t-1)/t}$ in the noisy case which implies immediately the results in the noiseless case given in Theorem 1.1.1. For $v \in \mathbb{R}^p$, we denote $v_{\max(k)}$ as v with all but the largest k entries in absolute value set to zero, and $v_{-\max(k)} = v - v_{\max(k)}$.

Let us consider the signal recovery model (1.1) in the setting where the observations contain noise and the signal is not exactly k -sparse. This is of significant interest for many applications. Two types of bounded noise settings,

$$z \in \mathcal{B}^{\ell_2}(\epsilon) \triangleq \{z : \|z\|_2 \leq \epsilon\} \quad \text{and} \quad z \in \mathcal{B}^{DS}(\epsilon) \triangleq \{z : \|Az\|_\infty \leq \epsilon\},$$

are of particular interest. The first bounded noise case was considered for example in Donoho et al. (2006). The second case is motivated by the *Dantzig Selector* procedure proposed in Candès and Tao (2007). Results on the Gaussian noise case, which is commonly studied in statistics, follow immediately. For notational convenience, we write δ and θ for RICs and ROCs of orders varying according to the scenarios.

Theorem 1.2.1. *Consider the signal recovery model (1.1) with $\|z\|_2 \leq \epsilon$. Suppose $\hat{\beta}^{\ell_2}$ is the minimizer of (1.2) with $\mathcal{B} = \mathcal{B}^{\ell_2}(\eta) = \{z : \|z\|_2 \leq \eta\}$ for some $\eta \geq \epsilon$.*

1. *If $\delta = \delta_k^A < 1/3$ for some $k \geq 2$, then*

$$\|\hat{\beta}^{\ell_2} - \beta\|_2 \leq \frac{\sqrt{2(1+\delta)}}{1-3\delta}(\epsilon + \eta) + \frac{\sqrt{2}(2\delta + \sqrt{(1-3\delta)\delta}) + (1-3\delta)}{1-3\delta} \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k}}. \quad (1.10)$$

2. If $\delta + \theta = \delta_k^A + \theta_{k,k}^A < 1$ for some $k \geq 1$, then

$$\|\hat{\beta}^{\ell_2} - \beta\|_2 \leq \frac{\sqrt{2(1+\delta)}}{1-\delta-\theta}(\epsilon + \eta) + \left(\frac{\sqrt{2\theta}}{1-\delta-\theta} + 1 \right) \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k}}. \quad (1.11)$$

3. If $\delta = \delta_{tk}^A < \sqrt{(t-1)/t}$ for some $t \geq 4/3$, then

$$\begin{aligned} \|\hat{\beta}^{\ell_2} - \beta\|_2 &\leq \frac{\sqrt{2(1+\delta)}}{1-\sqrt{t/(t-1)}\delta}(\epsilon + \eta) \\ &+ \left(\frac{\sqrt{2\delta} + \sqrt{t(\sqrt{(t-1)/t} - \delta)\delta}}{t(\sqrt{(t-1)/t} - \delta)} + 1 \right) \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k}}. \end{aligned} \quad (1.12)$$

Now consider the signal recovery model (1.1) with $\|A^T z\|_\infty \leq \epsilon$. Suppose $\hat{\beta}^{DS}$ is the minimizer of (1.2) with $\mathcal{B} = \mathcal{B}^{DS}(\eta) = \{z : \|A^T z\|_\infty \leq \eta\}$ for some $\eta \geq \epsilon$.

1. If $\delta = \delta_k^A < 1/3$ with $k \geq 2$, then

$$\|\hat{\beta}^{DS} - \beta\|_2 \leq \frac{\sqrt{2k}}{1-3\delta}(\epsilon + \eta) + \frac{\sqrt{2}(2\delta + \sqrt{(1-3\delta)\delta}) + (1-3\delta)}{1-3\delta} \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k}}. \quad (1.13)$$

2. If $\delta + \theta = \delta_k^A + \theta_{k,k}^A < 1$ for some $k \geq 1$, then

$$\|\hat{\beta} - \beta\|_2 \leq \frac{\sqrt{2k}}{1-\delta-\theta}(\epsilon + \eta) + \left(\frac{\sqrt{2\theta}}{1-\delta-\theta} + 1 \right) \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k}}. \quad (1.14)$$

3. If $\delta = \delta_{tk}^A < \sqrt{(t-1)/t}$ for some $t \geq 4/3$, then

$$\begin{aligned} \|\hat{\beta}^{DS} - \beta\|_2 &\leq \frac{\sqrt{2tk}}{1-\sqrt{t/(t-1)}\delta}(\epsilon + \eta) \\ &+ \left(\frac{\sqrt{2\delta} + \sqrt{t(\sqrt{(t-1)/t} - \delta)\delta}}{t(\sqrt{(t-1)/t} - \delta)} + 1 \right) \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k}}. \end{aligned} \quad (1.15)$$

Remark 1.2.1. The result for the noiseless case follows directly from Theorem 1.2.1. When β is exactly k -sparse and there is no noise, by setting $\eta = \epsilon = 0$ and by noting $\beta_{-\max(k)} = 0$, we have $\hat{\beta} = \beta$ from (1.12), where $\hat{\beta}$ is the minimizer of (1.2) with $\mathcal{B} = \{0\}$.

Remark 1.2.2. In the Part 1 of Theorems 1.1.1 and 1.2.1 (on $\delta_k^A < 1/3$), the case $k = 1$ is excluded because the RIC of order 1 cannot provide any sufficient condition for the exact recovery via the constrained ℓ_1 minimization in this case. Take, for example, $n = p - 1 \geq 1$. Let $A \in \mathbb{R}^{n \times p}$ with $A\beta = (\beta_1 - \beta_2, \beta_3, \beta_4, \dots, \beta_p)^T$ for any $\beta = (\beta_1, \beta_2, \beta_3, \dots, \beta_p)^T \in \mathbb{R}^p$. Then for all 1-sparse vectors β ,

$$\|A\beta\|_2^2 = \sum_{i=1}^p \beta_i^2 - 2\beta_1\beta_2 = \|\beta\|_2^2,$$

which implies the restricted isometry constant $\delta_1^A = 0$. However, $b = A\gamma = A\eta$ where $\gamma = (1, 0, \dots, 0)$ and $\eta = (0, -1, 0, \dots, 0)$ are both 1-sparse signals. Thus it is impossible to recover both of them exactly relying only on the information of (A, b) . In particular, the ℓ_1 minimization (1.2) with $\mathcal{B} = \{0\}$ cannot recover all 1-sparse signals. Since $\delta_1^A = 0$, the RIP cannot provide any sufficient condition in this case.

Remark 1.2.3. The condition $\delta_k^A + \theta_{k,k}^A < 1$ in Part 2 of Theorem 1.2.1 can be extended to a more general form,

$$\delta_a^A + C_{a,b,k} \theta_{a,b}^A < 1, \text{ where } C_{a,b,k} = \max \left\{ \frac{2k-a}{\sqrt{ab}}, \sqrt{\frac{2k-a}{a}} \right\}, 1 \leq a \leq k. \quad (1.16)$$

Theorem 1.2.2. *Let $y = A\beta$ where $\beta \in \mathbb{R}^p$ is a k -sparse vector. If the condition (1.16) holds, then the ℓ_1 norm minimizer $\hat{\beta}$ of (1.2) with $\mathcal{B} = \{0\}$ recovers β exactly.*

In the noisy case, we have the following theorem parallel to Part 2 of Theorem 1.2.1.

Theorem 1.2.3. Consider the signal recovery model (1.1) with $\|z\|_2 \leq \epsilon$. Let $\hat{\beta}$ be the minimizer of (1.2) with $\mathcal{B} = \{z \in \mathbb{R}^n : \|z\|_2 \leq \eta\}$ for some $\eta \geq \epsilon$. If $\delta + \theta = \delta_a^A + C_{a,b,k}\theta_{a,b}^A < 1$ for some positive integers a and b with $1 \leq a \leq k$, then

$$\|\hat{\beta} - \beta\|_2 \leq \frac{\sqrt{2(1+\delta)k/a}}{1-\delta-C_{a,b,k}\theta}(\epsilon+\eta) + 2\|\beta_{-\max(k)}\|_1 \left(\frac{\sqrt{2k}C_{a,b,k}\theta}{(1-\delta-C_{a,b,k}\theta)(2k-a)} + \frac{1}{\sqrt{k}} \right). \quad (1.17)$$

Similarly, consider the signal recovery model (1.1) with $\|A^T z\|_\infty \leq \epsilon$. Let $\hat{\beta}$ be the minimizer of (1.2) with $\mathcal{B} = \{z \in \mathbb{R}^n : \|A^T z\|_\infty \leq \eta\}$ for some $\eta \geq \epsilon$. If $\delta + \theta = \delta_a^A + C_{a,b,k}\theta_{a,b}^A < 1$ for some positive integers a and b with $1 \leq a \leq k$, then

$$\|\hat{\beta} - \beta\|_2 \leq \frac{\sqrt{2k}}{1-\delta-C_{a,b,k}\theta}(\epsilon+\eta) + 2\|\beta_{-\max(k)}\|_1 \left(\frac{\sqrt{2k}C_{a,b,k}\theta}{(1-\delta-C_{a,b,k}\theta)(2k-a)} + \frac{1}{\sqrt{k}} \right). \quad (1.18)$$

Remark 1.2.4. It should be noted that Part 3 of Theorem 1.2.1 also hold for $1 < t < 4/3$ with exactly the same proof. However the bound $\sqrt{(t-1)/t}$ is not sharp for $1 < t < 4/3$. See Section 1.4 for further discussions. The condition $t \geq 4/3$ is crucial for the “sharpness” results given in Theorem 1.2.4 at the end of this section.

The signal recovery model (1.1) with Gaussian noise is of particular interest in statistics and signal processing. The following results on the i.i.d. Gaussian noise case are immediate consequences of the above results on the bounded noise cases, since the Gaussian random variables are essentially bounded.

Proposition 1.2.1. Suppose the error vector $z \sim N_n(0, \sigma^2 I)$ in (1.1). Let $\hat{\beta}^{\ell_2}$ be the minimizer of (1.2) with $\mathcal{B} = \{z : \|z\|_2 \leq \sigma\sqrt{n+2\sqrt{n\log n}}\}$ and let $\hat{\beta}^{DS}$ be the minimizer of (1.2) with $\mathcal{B} = \{z : \|A^T z\|_\infty \leq 2\sigma\sqrt{\log p}\}$.

- If $\delta_k^A < 1/3$ for some $k \geq 2$, then with probability at least $1 - 1/n$,

$$\begin{aligned} \|\beta^{\ell_2} - \beta\|_2 &\leq \frac{2\sqrt{2(1+\delta)}}{1-3\delta} \sigma \sqrt{n + 2\sqrt{n \log n}} \\ &+ \frac{\sqrt{2}(2\delta + \sqrt{(1-3\delta)\delta}) + (1-3\delta) 2\|\beta_{-\max(k)}\|_1}{1-3\delta} \frac{1}{\sqrt{k}}, \end{aligned}$$

and with probability at least $1 - 1/\sqrt{\pi \log p}$,

$$\begin{aligned} \|\hat{\beta}^{DS} - \beta\|_2 &\leq \frac{4\sqrt{2}}{1-3\delta} \sigma \sqrt{k \log p} \\ &+ \frac{\sqrt{2}(2\delta + \sqrt{(1-3\delta)\delta}) + (1-3\delta) 2\|\beta_{-\max(k)}\|_1}{1-3\delta} \frac{1}{\sqrt{k}}. \end{aligned}$$

- If $\delta_k^A + \theta_{k,k}^A < 1$ for some $k \geq 1$, then with probability at least $1 - 1/n$,

$$\|\beta^{\ell_2} - \beta\|_2 \leq \frac{2\sqrt{2(1+\delta)}}{1-\delta-\theta} \sigma \sqrt{n + 2\sqrt{n \log n}} + \left(\frac{\sqrt{2}\theta}{1-\delta-\theta} + 1 \right) \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k}},$$

and with probability at least $1 - 1/\sqrt{\pi \log p}$,

$$\|\hat{\beta}^{DS} - \beta\|_2 \leq \frac{4\sqrt{2}}{1-\delta-\theta} \sigma \sqrt{k \log p} + \left(\frac{\sqrt{2}\theta}{1-\delta-\theta} + 1 \right) \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k}}.$$

- If $\delta_{tk}^A < \sqrt{(t-1)/t}$ for some $t \geq 4/3$, then with probability at least $1 - 1/n$,

$$\begin{aligned} \|\beta^{\ell_2} - \beta\|_2 &\leq \frac{2\sqrt{2(1+\delta)}}{1-\sqrt{t/(t-1)}\delta} \sigma \sqrt{n + 2\sqrt{n \log n}} \\ &+ \left(\frac{\sqrt{2}\delta + \sqrt{t(\sqrt{(t-1)/t} - \delta)\delta}}{t(\sqrt{(t-1)/t} - \delta)} + 1 \right) \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k}}, \end{aligned}$$

and with probability at least $1 - 1/\sqrt{\pi \log p}$,

$$\begin{aligned} \|\hat{\beta}^{DS} - \beta\|_2 &\leq \frac{4\sqrt{2t}}{1 - \sqrt{t/(t-1)}\delta} \sigma \sqrt{k \log p} \\ &+ \left(\frac{\sqrt{2}\delta + \sqrt{t(\sqrt{(t-1)/t} - \delta)\delta}}{t(\sqrt{(t-1)/t} - \delta)} + 1 \right) \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k}}. \end{aligned}$$

The oracle inequality approach was introduced by Donoho and Johnstone (1994) in the context of wavelet thresholding for signal denoising. It provides an effective way to study the performance of an estimation procedure by comparing it to that of an ideal estimator. In the context of compressed sensing, oracle inequalities have been given in Cai et al. (2010d), Candès and Tao (2007) and Candès and Plan (2011) under various settings. Proposition 1.2.2 below provides oracle inequalities for compressed sensing with Gaussian noise under the conditions $\delta_k^A < 1/3$, $\delta_k^A + \theta_{k,k}^A < 1$ or $\delta_{tk}^A < \sqrt{(t-1)/t}$ when $t \geq 4/3$.

Proposition 1.2.2. *Given (1.1), suppose the error vector $z \sim N_n(0, \sigma^2 I)$, β is k -sparse. Let $\hat{\beta}^{DS}$ be the minimizer of (1.2) with $\mathcal{B} = \{z : \|A^T z\|_\infty \leq 4\sigma\sqrt{\log p}\}$.*

- If $\delta_k^A < 1/3$ for some $k \geq 2$, then with probability at least $1 - 1/\sqrt{\pi \log p}$,

$$\|\hat{\beta}^{DS} - \beta\|_2^2 \leq \frac{256}{(1 - 3\delta_k^A)^2} \log p \sum_i \min(\beta_i^2, \sigma^2). \quad (1.19)$$

- If $\delta_k^A + \theta_{k,k}^A < 1/3$ for some $k \geq 1$, then with probability at least $1 - 1/\sqrt{\pi \log p}$,

$$\|\hat{\beta}^{DS} - \beta\|_2^2 \leq \frac{256}{(1 - \delta_k^A - \theta_{k,k}^A)^2} \log p \sum_i \min(\beta_i^2, \sigma^2). \quad (1.20)$$

- If $\delta_{tk}^A < \sqrt{(t-1)/t}$ for some $t \geq 4/3$, then with probability at least $1 -$

$1/\sqrt{\pi \log p}$,

$$\|\hat{\beta}^{DS} - \beta\|_2^2 \leq \frac{256t}{(1 - \sqrt{t/(t-1)}\delta_{tk}^A)^2} \log p \sum_i \min(\beta_i^2, \sigma^2). \quad (1.21)$$

We now turn to show the sharpness of the conditions $\delta_k^A < 1/3$, $\delta_k^A + \theta_{k,k}^A < 1$ and $\delta_{tk}^A < \sqrt{(t-1)/t}$ for the exact recovery in the noiseless case and stable recovery in the noisy case. It should be noted that the result in the special case $t = 2$ was shown in Davies and Gribonval (2009).

Theorem 1.2.4.

1. For all $2 \leq k \leq p/2$, there exists a sensing matrix A satisfying $\delta_k^A = 1/3$,
2. For all $2 \leq k \leq p/2$, there exists a sensing matrix A satisfying $\delta_k^A + \theta_{k,k}^A = 1$,
3. Let $t \geq 4/3$. For all $\epsilon > 0$ and $k \geq 5/\epsilon$, there exists a sensing matrix A satisfying $\delta_{tk}^A < \sqrt{\frac{t-1}{t}} + \epsilon$,

and in any of the three scenarios above, there also exists some k -sparse vector β_0 such that

- in the noiseless case, i.e. $y = A\beta_0$, the ℓ_1 minimization method (1.2) with $\mathcal{B} = \{0\}$ fail to exactly recover the k -sparse vector β_0 , i.e. $\hat{\beta} \neq \beta_0$, where $\hat{\beta}$ is the solution to (1.2).
- in the noisy case, i.e. $y = A\beta_0 + z$, for all constraints \mathcal{B}_z (may depends on z), the ℓ_1 minimization method (1.2) fails to stably recover the k -sparse vector β_0 , i.e. $\hat{\beta} \not\rightarrow \beta$ as $z \rightarrow 0$, where $\hat{\beta}$ is the solution to (1.2).

Remark 1.2.5. Similarly as Theorems 1.1.1 and 1.2.1, there is a more general form of Part 2 of Theorem 1.2.4 on Condition $\delta_k^A + \theta_{k,k}^A < 1$, which is stated below.

Theorem 1.2.5. *Let $1 \leq k \leq p/2$, $1 \leq a \leq k$, and $b \geq 1$. Let $C_{a,b,k}$ be defined as (1.16). Then there exists a sensing matrix $A \in \mathbb{R}^{n \times p}$ such that $\delta_a^A + C_{a,b,k}\theta_{a,b}^A = 1$ and for some k -sparse signals $\beta_0 \in \mathbb{R}^p$ such that the conclusion in Theorem 1.2.4 still holds.*

1.3 Affine Rank Minimization

We consider the affine rank minimization problem (1.3) in this section. As mentioned in the introduction, this problem is closely related to compressed sensing. The close connections between compressed sensing and ARMP have been studied in Oymak, et al. Oymak et al. (2011). We shall present here the analogous results on affine rank minimization without detailed proofs.

For a matrix $X \in \mathbb{R}^{m \times n}$ (without loss of generality, assume that $m \leq n$) with the singular value decomposition $X = \sum_{i=1}^m a_i u_i v_i^T$ where the singular values a_i are in descending order, we define $X_{\max(r)} = \sum_{i=1}^r a_i u_i v_i^T$ and $X_{-\max(r)} = \sum_{i=r+1}^m a_i u_i v_i^T$. We should also note that the nuclear norm $\|\cdot\|_*$ of a matrix equals the sum of the singular values, and the spectral norm $\|\cdot\|$ of a matrix equals its largest singular value. Their roles are similar to those of ℓ_1 norm and ℓ_∞ norm in the vector case, respectively. For a linear operator $\mathcal{M} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^q$, its dual operator is denoted by $\mathcal{M}^* : \mathbb{R}^q \rightarrow \mathbb{R}^{m \times n}$.

Similarly as in compressed sensing, we first consider the matrix recovery model (1.3) in the case where the error vector z is in bounded sets: $\|z\|_2 \leq \epsilon$ and $\|\mathcal{M}^*(z)\| \leq \epsilon$. The corresponding nuclear norm minimization methods are given by (1.4) with $\mathcal{B} = \mathcal{B}^{\ell_2}(\eta)$ and $\mathcal{B} = \mathcal{B}^{DS}(\eta)$ respectively, where

$$\mathcal{B}^{\ell_2}(\eta) = \{z : \|z\|_2 \leq \eta\}, \quad (1.22)$$

$$\mathcal{B}^{DS}(\eta) = \{z : \|\mathcal{M}^*(z)\| \leq \eta\}. \quad (1.23)$$

Proposition 1.3.1. Consider ARMP (1.3) with $\|z\|_2 \leq \epsilon$. Let $X_*^{\ell_2}$ be the minimizer of (1.4) with $\mathcal{B} = \mathcal{B}^{\ell_2}(\eta)$ defined in (1.22) for some $\eta \geq \epsilon$.

1. If $\delta = \delta_r^{\mathcal{M}} < 1/3$ for some $r \geq 2$, then

$$\|X_*^{\ell_2} - X\|_F \leq \frac{\sqrt{2(1+\delta)}}{1-3\delta}(\epsilon+\eta) + \frac{\sqrt{2}(2\delta + \sqrt{(1-3\delta)\delta}) + (1-3\delta)}{1-3\delta} \frac{2\|X_{-\max(r)}\|_*}{\sqrt{r}}. \quad (1.24)$$

2. If $\delta + \theta = \delta_r^{\mathcal{M}} + \theta_{r,r}^{\mathcal{M}} < 1$ for some $r \geq 1$, then

$$\|X_*^{\ell_2} - X\|_F \leq \frac{\sqrt{2(1+\delta)}}{1-\delta-\theta}(\epsilon+\eta) + \left(\frac{\sqrt{2}\theta}{1-\delta-\theta} + 1 \right) \frac{2\|X_{-\max(r)}\|_*}{\sqrt{r}}. \quad (1.25)$$

3. If $\delta = \delta_{tr}^{\mathcal{M}} < \sqrt{(t-1)/t}$ with $t \geq 4/3$, then

$$\begin{aligned} \|X_*^{\ell_2} - X\|_F &\leq \frac{\sqrt{2(1+\delta)}}{1-\sqrt{t/(t-1)}\delta}(\epsilon+\eta) \\ &+ \left(\frac{\sqrt{2}\delta + \sqrt{t(\sqrt{(t-1)/t}-\delta)\delta}}{t(\sqrt{(t-1)/t}-\delta)} + 1 \right) \frac{2\|X_{-\max(r)}\|_*}{\sqrt{r}}. \end{aligned} \quad (1.26)$$

Similarly, consider ARMP (1.3) with z satisfying $\|\mathcal{M}^*(z)\| \leq \epsilon$. Let X_*^{DS} be the minimizer of (1.4) with $\mathcal{M} = \mathcal{B}^{DS}(\eta)$ defined in (1.23) for some $\eta \geq \epsilon$.

1. If $\delta = \delta_r^{\mathcal{M}} < 1/3$ with $r \geq 2$, then

$$\|X_*^{DS} - X\|_F \leq \frac{\sqrt{2r}}{1-3\delta}(\epsilon+\eta) + \frac{\sqrt{2}(2\delta + \sqrt{(1-3\delta)\delta}) + (1-3\delta)}{1-3\delta} \frac{2\|X_{-\max(r)}\|_*}{\sqrt{r}}. \quad (1.27)$$

2. If $\delta + \theta = \delta_r^{\mathcal{M}} + \theta_{r,r}^{\mathcal{M}} < 1$ for some $r \geq 1$, then

$$\|X_*^{DS} - X\|_F \leq \frac{\sqrt{2r}}{1-\delta-\theta}(\epsilon+\eta) + \left(\frac{\sqrt{2}\theta}{1-\delta-\theta} + 1 \right) \frac{2\|X_{-\max(r)}\|_*}{\sqrt{r}}. \quad (1.28)$$

3. If $\delta = \delta_{tr}^{\mathcal{M}} < \sqrt{(t-1)/t}$ with $t \geq 4/3$, then

$$\begin{aligned} \|X_*^{DS} - X\|_F &\leq \frac{\sqrt{2tr}}{1 - \sqrt{t/(t-1)}\delta}(\epsilon + \eta) \\ &+ \left(\frac{\sqrt{2}\delta + \sqrt{t(\sqrt{(t-1)/t} - \delta)\delta}}{t(\sqrt{(t-1)/t} - \delta)} + 1 \right) \frac{2\|X_{-\max(r)}\|_1}{\sqrt{r}}. \end{aligned} \quad (1.29)$$

In the special noiseless case where $z = 0$, it can be seen from either of these two inequalities above that all matrices X with rank at most r can be exactly recovered provided that $\delta_r^{\mathcal{M}} < 1/3$, $\delta_r^{\mathcal{M}} + \theta_{r,r}^{\mathcal{M}} < 1$ or $\delta_{tr}^{\mathcal{M}} < \sqrt{(t-1)/t}$, for some $t \geq 4/3$.

In the matrix recover model (1.3) with Gaussian noise, oracle inequalities can also be developed under conditions $\delta_r^{\mathcal{M}} < 1/3$, $\delta_r^{\mathcal{M}} + \theta_{r,r}^{\mathcal{M}} < 1$, or $\delta_{tr}^{\mathcal{M}} < \sqrt{(t-1)/t}$ when $t \geq 4/3$.

Proposition 1.3.2. *Given ARMP (1.3), suppose the error vector $z \sim N_q(0, \sigma^2 I)$, $\text{rank}(\beta) \leq r$. Let \hat{X}^{DS} be the minimizer of (1.2) with $\mathcal{B} = \{z : \|\mathcal{M}^*z\| \leq \lambda = 8\sigma\sqrt{2\log(12)\max(m,n)}\}$.*

- If $\delta_r^{\mathcal{M}} < 1/3$ for some $r \geq 2$, then with probability at least $1 - e^{-c\max(m,n)}$,

$$\|X_*^{DS} - X\|_F^2 \leq \frac{2^{11}\log(12)}{(1 - 3\delta_r^{\mathcal{M}})^2} \sum_i \min(\sigma_i^2(X), \max(m,n)\sigma^2). \quad (1.30)$$

- If $\delta_r^{\mathcal{M}} + \theta_{r,r}^{\mathcal{M}} < 1$ for some $r \geq 1$, then with probability at least $1 - e^{-c\max(m,n)}$,

$$\|X_*^{DS} - X\|_F^2 \leq \frac{2^{11}\log(12)}{(1 - \delta_r^{\mathcal{M}} - \theta_{r,r}^{\mathcal{M}})^2} \sum_i \min(\sigma_i^2(X), \max(m,n)\sigma^2). \quad (1.31)$$

- If $\delta_{tr}^{\mathcal{M}} < \sqrt{(t-1)/t}$ for some $t \geq 4/3$, then with probability at least $1 -$

$$e^{-c \max(m,n)},$$

$$\|X_*^{DS} - X\|_F^2 \leq \frac{2^{11} \log(12)t}{(1 - \sqrt{t/(t-1)})\delta_{tr}^{\mathcal{M}})^2} \sum_i \min(\sigma_i^2(X), \max(m, n)\sigma^2). \quad (1.32)$$

Here $c > 0$ is an absolute constant, and $\sigma_i(X), i = 1, \dots, \min(m, n)$ are the singular values of X .

The following result shows that the conditions $\delta_r^{\mathcal{M}} < 1/3$, $\delta_r^{\mathcal{M}} + \theta_{r,r}^{\mathcal{M}} < 1$, $\delta_{tr}^{\mathcal{M}} < \sqrt{(t-1)/t}$ with $t \geq 4/3$ are sharp. These results together establish the optimal bounds on $\delta_r^{\mathcal{M}}$, $\delta_r^{\mathcal{M}} + \theta_{r,r}^{\mathcal{M}}$ and $\delta_{tr}^{\mathcal{M}}$ ($t \geq 4/3$) for the exact recovery in the noiseless case.

Proposition 1.3.3.

1. For all $2 \leq r \leq p/2$, there exists a linear map \mathcal{M} satisfying $\delta_r^{\mathcal{M}} = 1/3$,
2. For all $2 \leq k \leq p/2$, there exists a linear map \mathcal{M} satisfying $\delta_r^{\mathcal{M}} + \theta_{r,r}^{\mathcal{M}} = 1$,
3. Let $t \geq 4/3$. For all $\epsilon > 0$ and $r \geq 5/\epsilon$, there exists a linear map A satisfying

$$\delta_{tr}^{\mathcal{M}} < \sqrt{\frac{t-1}{t}} + \epsilon,$$

and in any of the three scenarios above, there also exists some matrix X_0 of rank at most r such that

- in the noiseless case, i.e. $b = \mathcal{M}(X_0)$, the nuclear norm minimization method (1.4) with $\mathcal{B} = \{0\}$ fails to exactly recover X_0 , i.e. $X_* \neq X_0$, where X_* is the solution to (1.4).
- in the noisy case, i.e. $b = \mathcal{M}(X_0) + z$, for all constraints \mathcal{B}_z (may depends on z), the nuclear norm minimization method (1.4) fails to stably recover X_0 , i.e. $X_* \not\rightarrow X_0$ as $z \rightarrow 0$, where X_* is the solution to (1.4) with $\mathcal{B} = \mathcal{B}_z$.

1.4 Discussions

We shall focus the discussions in this section exclusively on compressed sensing as the results on affine rank minimization is analogous. In Section 1.2, we have established the sharp RIP condition on different orders of RICs,

$$\delta_k^A < 1/3$$

$$\delta_{tk}^A < \sqrt{\frac{t-1}{t}} \quad \text{for some } t \geq \frac{4}{3},$$

for the recovery of k -sparse signals in compressed sensing. For a general $t > 0$, denote the sharp bound for δ_{tk}^A as $\delta_*(t)$. Then

$$\delta_*(1) = 1/3 \quad \text{and} \quad \delta_*(t) = \sqrt{(t-1)/t}, \quad t \geq 4/3.$$

A natural question is: What is the value of $\delta_*(t)$ for $t < 4/3$ and $t \neq 1$? That is, what is the sharp bound for δ_{tk}^A when $t < 4/3$ and $t \neq 1$? We have the following partial answer to the question.

Proposition 1.4.1. *Let $y = A\beta$ where $\beta \in \mathbb{R}^p$ is k -sparse. Suppose $0 < t < 1$ and $tk \geq 0$ to be an integer*

- *When tk is even and $\delta_{tk}^A < \frac{t}{4-t}$, the ℓ_1 minimization (1.2) with $\mathcal{B} = \{0\}$ recovers β exactly.*
- *When tk is odd and $\delta_{tk}^A < \frac{\sqrt{t^2-1/k^2}}{4-2t+\sqrt{t^2-1/k^2}}$, the ℓ_1 minimization (1.2) with $\mathcal{B} = \{0\}$ recovers β exactly.*

In addition, the following result shows that $\delta_*(t) \leq \frac{t}{4-t}$ for all $0 < t < 4/3$. In particular, when $t = 1$, the upper bound $t/(4-t)$ coincides with the true sharp bound $1/3$.

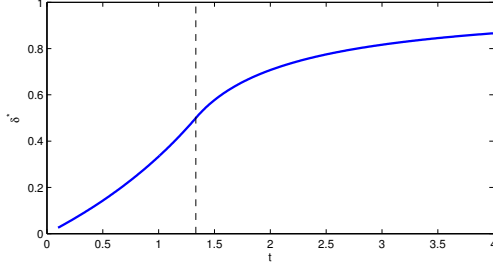


Figure 1.2: Plot of δ_* as a function of t . The dotted line is $t = 4/3$.

Proposition 1.4.2. *For $0 < t < 4/3$, $\epsilon > 0$ and any integer $k \geq 1$, $\delta_{tk}^A < \frac{t}{4-t} + \epsilon$ is not sufficient for the exact recovery. Specifically, there exists a matrix A with $\delta_{tk}^A = \frac{t}{4-t}$ and a k -sparse vector β_0 such that $\hat{\beta} \neq \beta_0$, where $\hat{\beta}$ is the minimizer of (1.2) with $\mathcal{B} = \{0\}$.*

Propositions 1.4.1 and 1.4.2 together show that $\delta_*(t) = \frac{t}{4-t}$ when tk is even and $0 < t < 1$. We are not able to provide a complete answer for $\delta_*(t)$ when $0 < t < 4/3$. We conjecture that $\delta_*(t) = \frac{t}{4-t}$ for all $0 < t < 4/3$. Figure 1.2 plots $\delta_*(t)$ as a function of t based on this conjecture for the interval $(0, 4/3)$.

Our results show that exact recovery of k -sparse signals in the noiseless case is guaranteed if $\delta_{tk}^A < \sqrt{(t-1)/t}$ for some $t \geq 4/3$. It is then natural to ask the question: Among all these RIP conditions $\delta_{tk}^A < \delta_*(t)$, which one is easiest to be satisfied? There is no general answer to this question as no condition is strictly weaker or stronger than the others. It is however interesting to consider special random measurement matrices $A = (A_{ij})_{n \times p}$ where

$$A_{ij} \sim \mathcal{N}(0, 1/n), \quad A_{ij} \sim \begin{cases} 1/\sqrt{n} & \text{w.p. } 1/2 \\ -1/\sqrt{n} & \text{w.p. } 1/2 \end{cases}, \quad \text{or} \quad A_{ij} \sim \begin{cases} \sqrt{3/n} & \text{w.p. } 1/6 \\ 0 & \text{w.p. } 1/2 \\ -\sqrt{3/n} & \text{w.p. } 1/6 \end{cases}.$$

Baraniuk et al. (2008) provides a bound on RICs for a set of random matrices from

concentration of measure. For these random measurement matrices, Theorem 5.2 of Baraniuk et al. (2008) shows that for positive integer $m < n$ and $0 < \lambda < 1$,

$$P(\delta_m^A < \lambda) \geq 1 - 2 \left(\frac{12ep}{m\lambda} \right)^m \exp(-n(\lambda^2/16 - \lambda^3/48)). \quad (1.33)$$

Hence, for $t \geq 4/3$,

$$P(\delta_{tk}^A < \sqrt{(t-1)/t}) \geq 1 - 2 \exp \left(tk \left(\log(12e/\sqrt{t(t-1)}) + \log(p/k) \right) - n \left(\frac{t-1}{16t} - \frac{(t-1)^{3/2}}{48t^{3/2}} \right) \right).$$

For $0 < t < 4/3$, using the conjectured value $\delta_*(t) = \frac{t}{4-t}$, we have

$$P(\delta_{tk}^A < t/(4-t)) \geq 1 - 2 \exp \left(tk(\log(12(4-t)e/t^2) + \log(p/k)) - n \left(\frac{t^2}{16(4-t)^2} - \frac{t^3}{48(4-t)^3} \right) \right).$$

It is easy to see when p, k , and $p/k \rightarrow \infty$, the lower bound of n to ensure $\delta_{tk}^A < t/(4-t)$ or $\delta_{tk}^A < \sqrt{(t-1)/t}$ to hold in high probability is $n \geq k \log(p/k)n^*(t)$, where

$$n^* \triangleq \begin{cases} t / \left(\frac{t^2}{16(4-t)^2} - \frac{t^3}{48(4-t)^3} \right) & t < 4/3; \\ t / \left(\frac{t-1}{16t} - \frac{(t-1)^{3/2}}{48t^{3/2}} \right), & t \geq 4/3. \end{cases}$$

For the plot of $n^*(t)$, see Figure 1.3. $n^*(t)$ has minimum 83.2 when $t = 1.85$. Moreover, among integer t , $t = 2$ can also provide a near-optimal minimum: $n^*(2) = 83.7$.

We should note that the above analysis is based on the bound given in (1.33) which itself can be possibly improved.

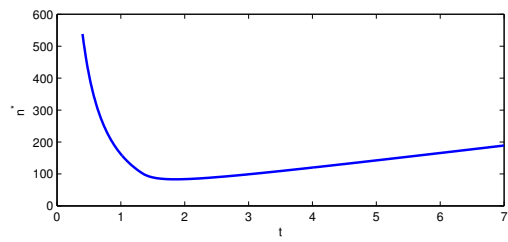


Figure 1.3: Plot of n^* as a function of t .

ROP: Matrix Recovery via Rank-one Projections

2.1 Introduction

Accurate recovery of low-rank matrices has a wide range of applications, including quantum state tomography (Alquier et al. 2013, Gross et al. 2010), face recognition (Basri and Jacobs 2003, Candès et al. 2011), recommender systems (Koren et al., 2009), and linear system identification and control (Recht et al., 2010). For example, a key step in reconstructing the quantum states in low-rank quantum tomography is the estimation of a low-rank matrix based on Pauli measurements (Gross et al. 2010, Wang 2013). And phase retrieval, a problem which arises in a range of signal and image processing applications including X-ray crystallography, astronomical imaging, and diffraction imaging, can be reformulated as a low-rank matrix recovery problem Candès et al. (2013); Candès et al. (2011). See Recht et al. (2010) and Candès and Plan (2011) for further references and discussions.

Motivated by these applications, low-rank matrix estimation based on a small number of measurements has drawn much recent attention in several fields, including statistics, electrical engineering, applied mathematics, and computer science. For example, Candès and Recht (2009), Candès and Tao (2010) and Recht Recht (2011) considered the exact recovery of a low-rank matrix based on a subset of uniformly

sampled entries. Negahban and Wainwright (2011) investigated matrix completion under a row/column weighted random sampling scheme. Recht et al. (2010), Candès and Plan (2011), and Cai and Zhang (2013b,a, 2014b) studied matrix recovery based on a small number of linear measurements in the framework of restricted isometry property (RIP), and Koltchinskii et al. (2011) proposed the penalized nuclear norm minimization method and derived a general sharp oracle inequality under the condition of restricted isometry in expectation.

The basic model for low-rank matrix recovery can be written as

$$y = \mathcal{X}(A) + z, \quad (2.1)$$

where $\mathcal{X} : \mathbb{R}^{p_1 \times p_2} \rightarrow \mathbb{R}^n$ is a linear map, $A \in \mathbb{R}^{p_1 \times p_2}$ is an unknown low-rank matrix, and z is a noise vector. The goal is to recover the low-rank matrix A based on the measurements (\mathcal{X}, y) . The linear map \mathcal{X} can be equivalently specified by n $p_1 \times p_2$ measurement matrices X_1, \dots, X_n with

$$\mathcal{X}(A) = (\langle X_1, A \rangle, \langle X_2, A \rangle, \dots, \langle X_n, A \rangle)^\top, \quad (2.2)$$

where the inner product of two matrices of the same dimensions is defined as $\langle X, Y \rangle = \sum_{i,j} X_{ij} Y_{ij}$. Since $\langle X, Y \rangle = \text{Trace}(X^\top Y)$, (2.1) is also known as trace regression.

A common approach to low-rank matrix recovery is the constrained nuclear norm minimization method which estimates A by

$$\hat{A} = \arg \min_M \{\|M\|_* : y - \mathcal{X}(M) \in \mathcal{Z}\}. \quad (2.3)$$

Here $\|X\|_*$ is the nuclear norm of the matrix X which is defined to be the sum of its singular values, and \mathcal{Z} is a bounded set determined by the noise structure. For

example, $\mathcal{Z} = \{0\}$ in the noiseless case and \mathcal{Z} is the feasible set of the error vector z in the case of bounded noise. This constrained nuclear norm minimization method has been well studied. See, for example, (Recht et al. 2010, Candès and Plan 2011, Oymak and Hassibi 2010, Cai and Zhang 2013b,a, 2014b).

Two random design models for low-rank matrix recovery have been particularly well studied in the literature. One is the so-called ‘‘Gaussian ensemble’’ (Recht et al. 2010, Candès and Plan 2011), where the measurement matrices X_1, \dots, X_n are random matrices with i.i.d. Gaussian entries. By exploiting the low-dimensional structure, the number of linear measurements can be far smaller than the number of entries in the matrix to ensure stable recovery. It has been shown that a matrix A of rank r can be stably recovered by nuclear norm minimization with high probability, provided that $n \gtrsim r(p_1 + p_2)$ (Candès and Plan, 2011). One major disadvantage of the Gaussian ensemble design is that it requires $O(np_1p_2)$ bytes of storage space for \mathcal{X} , which can be excessively large for the recovery of large matrices. For example, at least 45 TB of space is need to store the measurement matrices M_i in order to ensure accurate reconstruction of 10000×10000 matrices of rank 10. (See more discussion in Section 2.5.) Another popular design is the ‘‘matrix completion’’ model (Candès and Recht 2009, Candès and Tao 2010, Recht 2011), under which the individual entries of the matrix A are observed at randomly selected positions. In terms of the measurement matrices X_i in (2.2), this can be interpreted as

$$\mathcal{X}(A) = (\langle e_{i_1} e_{j_1}^\top, A \rangle, \langle e_{i_2} e_{j_2}^\top, A \rangle, \dots, \langle e_{i_n} e_{j_n}^\top, A \rangle)^\top \quad (2.4)$$

where $e_i = (0, \dots, 0, \overbrace{1}^{\text{ith}}, 0, \dots, 0)$ is the i th standard basis vector, and i_1, \dots, i_n and j_1, \dots, j_n are randomly and uniformly drawn with replacement from $\{1, \dots, p_1\}$ and $\{1, \dots, p_2\}$, respectively. However, as pointed out in Candès and Recht (2009),

Recht (2011), additional structural assumptions, which are not intuitive and difficult to check, on the unknown matrix A are needed in order to ensure stable recovery under the matrix completion model. For example, it is impossible to recover spiked matrices under the matrix completion model. This can be easily seen from a simple example where the matrix A has only one non-zero row. In this case, although the matrix is only of rank one, it is not recoverable under the matrix completion model unless all the elements on the non-zero row are observed.

In this chapter we introduce a “*Rank-One Projection*” (*ROP*) model for low-rank matrix recovery and propose a constrained nuclear norm minimization method for this model. Under the ROP model, we observe

$$y_i = (\beta^{(i)})^\top A \gamma^{(i)} + z_i, \quad i = 1, \dots, n \quad (2.5)$$

where $\beta^{(i)}$ and $\gamma^{(i)}$ are random vectors with entries independently drawn from some distribution \mathcal{P} , and z_i are random errors. In terms of the linear map $\mathcal{X} : \mathbb{R}^{p_1 \times p_2} \rightarrow \mathbb{R}^n$ in (2.1), it can be defined as

$$[\mathcal{X}(A)]_i = (\beta^{(i)})^\top A \gamma^{(i)}, \quad i = 1, \dots, n. \quad (2.6)$$

Since the measurement matrices $X_i = \beta^{(i)}(\gamma^{(i)})^\top$ are of rank-one, we call the model (2.5) a “*Rank-One Projection*” (*ROP*) model. It is easy to see that the storage for the measurement vectors in the ROP model (2.5) is $O(n(p_1 + p_2))$ bytes which is significantly smaller than $O(np_1p_2)$ bytes required for the Gaussian ensemble.

We first establish a sufficient identifiability condition in Section 2.2 by considering the problem of exact recovery of low-rank matrices in the noiseless case. It is shown that, with high probability, ROP with $n \gtrsim r(p_1 + p_2)$ random projections is sufficient to ensure exact recovery of all rank r matrices through the constrained nuclear norm

minimization. The required number of measurements $O(r(p_1 + p_2))$ is rate optimal for any linear measurement model since a rank r matrix $A \in \mathbb{R}^{p_1+p_2}$ has the degree of freedom $r(p_1 + p_2 - r)$. The Gaussian noise case is of particular interest in statistics. We propose a new constrained nuclear norm minimization estimator and investigate its theoretical and numerical properties in the Gaussian noise case. Both upper and lower bounds for the estimation accuracy under the Frobenius norm loss are obtained. The estimator is shown to be rate-optimal when the number of rank-one projections satisfies either $n \gtrsim (p_1 + p_2) \log(p_1 + p_2)$ or $n \sim r(p_1 + p_2)$. The lower bound also shows that if the number of measurements $n < r \max(p_1, p_2)$, then no estimator can recover rank- r matrices consistently. The general case where the matrix A is only approximately low-rank is also considered. The results show that the proposed estimator is adaptive to the rank r and robust against small perturbations. Extensions to the sub-Gaussian design and sub-Gaussian noise distribution are also considered.

The ROP model can be further simplified by taking $\beta^{(i)} = \gamma^{(i)}$ if the low-rank matrix A is known to be symmetric. This is the case in many applications, including low-dimensional Euclidean embedding (Trosset 2000, Recht et al. 2010), phase retrieval (Candès et al. 2013, Candès et al. 2011), and covariance matrix estimation (Chen et al. 2013, Cai et al. 2013b,a). In such a setting, the ROP design can be simplified to symmetric rank-one projections (SROP)

$$[\mathcal{X}(A)]_i = (\beta^{(i)})^\top A \beta^{(i)}.$$

We will show that the results for the general ROP model continue to hold for the SROP model when A is known to be symmetric. Recovery of symmetric positive definite matrices in the noiseless and ℓ_1 -bounded noise settings has also been considered in a recent paper by Chen et al. (2013) which was posted on arXiv at the time of the writing of the present chapter. Their results and techniques for symmetric positive

definite matrices are not applicable to the recovery of general low-rank matrices. See Section 2.6 for more discussions.

The techniques and main results developed in the chapter also have implications to other related statistical problems. In particular, the results imply that it is possible to accurately estimate a spiked covariance matrix based only on one-dimensional projections. Spiked covariance matrix model has been well studied in the context of principal component analysis (PCA) based on i.i.d. data where one observes p -dimensional vectors $X^{(1)}, \dots, X^{(n)} \stackrel{iid}{\sim} N(0, \Sigma)$ with $\Sigma = I_p + \Sigma_0$ and Σ_0 being low-rank (Johnstone 2001, Birnbaum et al. 2013, Cai et al. 2013b,a). This covariance structure and its variations have been used in many applications including signal processing, financial econometrics, chemometrics, and population genetics. See, for example, Fan et al. (2008), Nadler (2010), Patterson and Reich (2006), Price et al. (2006), Wax and Kailath (1985). Suppose that the random vectors $X^{(1)}, \dots, X^{(n)}$ are not directly observable. Instead, we observe only one-dimensional random projections of $X^{(i)}$,

$$\xi_i = \langle \beta^{(i)}, X^{(i)} \rangle, \quad i = 1, \dots, n,$$

where $\beta^{(i)} \stackrel{iid}{\sim} N(0, I_p)$. It is somewhat surprising that it is still possible to accurately estimate the spiked covariance matrix Σ based only on the one-dimensional projections $\{\xi_i : i = 1, \dots, n\}$. This covariance matrix recovery problem is also related to the recent literature on covariance sketching (Dasarathy et al., 2012, 2013), which aims to recover a symmetric matrix A (or a general rectangular matrix B) from low-dimensional projections of the form $X^\top AX$ (or $X^\top BY$). See Section 2.4 for further discussions.

The proposed methods can be efficiently implemented via convex programming. A simulation study is carried out to investigate the numerical performance of the proposed nuclear norm minimization estimators. The numerical results indicate that

ROP with $n \geq 5r \max(p_1, p_2)$ random projections is sufficient to ensure the exact recovery of rank r matrices through constrained nuclear norm minimization and show that the procedure is robust against small perturbations, which confirm the theoretical results developed in the chapter. The proposed estimator outperforms two other alternative procedures numerically in the noisy case. In addition, the proposed method is illustrated through an image compression example.

The rest of this chapter is organized as follows. In Section 2.2, after introducing basic notation and definitions, we consider exact recovery of low-rank matrices in the noiseless case and establish a sufficient identifiability condition. A constrained nuclear norm minimization estimator is introduced for the Gaussian noise case. Both upper and lower bounds are obtained for estimation under the Frobenius norm loss. Section 2.3 considers extensions to sub-Gaussian design and sub-Gaussian noise distributions. An application to estimation of spiked covariance matrices based on one-dimensional projections is discussed in detail in Section 2.4. Section 2.5 investigates the numerical performance of the proposed procedure through a simulation study and an image compression example. A brief discussion is given in Section 2.6. The main results are proved in the Appendix (Chapter A.2).

2.2 Matrix Recovery under Gaussian Noise

In this section, we first establish an identifiability condition for the ROP model by considering exact recovery in the noiseless case, and then focus on low-rank matrix recovery in the Gaussian noise case.

We begin with the basic notation and definitions. For a vector $\beta \in \mathbb{R}^n$, we use $\|\beta\|_q = \sqrt[q]{\sum_{i=1}^n |\beta_i|^q}$ to define its vector q -norm. For a matrix $X \in \mathbb{R}^{p_1 \times p_2}$, the Frobenius norm is $\|X\|_F = \sqrt{\sum_{i=1}^{p_1} \sum_{j=1}^{p_2} X_{ij}^2}$ and the spectral norm $\|\cdot\|$ is $\|X\| = \sup_{\|\beta\|_2 \leq 1} \|X\beta\|_2$. For a linear map $\mathcal{X} = (X_1, \dots, X_n)$ from $\mathbb{R}^{p_1 \times p_2}$ to \mathbb{R}^n given by

(2.2), its dual operator $\mathcal{X}^* : \mathbb{R}^n \rightarrow \mathbb{R}^{p_1 \times p_2}$ is defined as $\mathcal{X}^*(z) = \sum_{i=1}^n z_i X_i$. For a matrix $X \in \mathbb{R}^{p_1 \times p_2}$, let $X = \sum_i a_i u_i v_i^\top$ be the singular value decomposition of X with the singular values $a_1 \geq a_2 \geq \dots \geq 0$. We define $X_{\max(r)} = \sum_{i=1}^r a_i u_i v_i^\top$ and $X_{-\max(r)} = X - X_{\max(r)} = \sum_{i \geq r+1} a_i u_i v_i^\top$. For any two sequences $\{a_n\}$ and $\{b_n\}$ of positive numbers, denote by $a_n \gtrsim b_n$ when $a_n \geq C b_n$ for some uniform constant C and denote by $a_n \sim b_n$ if $a_n \gtrsim b_n$ and $b_n \gtrsim a_n$.

We use the phrase ‘‘rank- r matrices’’ to refer to matrices of rank at most r and denote by \mathbb{S}^p the set of all $p \times p$ symmetric matrices. A linear map $\mathcal{X} : \mathbb{R}^{p_1 \times p_2} \rightarrow \mathbb{R}^n$ is called ROP from distribution \mathcal{P} if \mathcal{X} is defined as in (2.6) with all the entries of $\beta^{(i)}$ and $\gamma^{(i)}$ independently drawn from the distribution \mathcal{P} .

2.2.1 RUB, Identifiability, and Exact Recovery in the Noiseless Case

An important step towards understanding the constrained nuclear norm minimization is the study of exact recovery of low-rank matrices in the noiseless case which also leads to a sufficient identifiability condition. A widely used framework in the low-rank matrix recovery literature is the Restricted Isometry Property (RIP) in the matrix setting. See Recht et al. (2010), Candès and Plan (2011), Rohde and Tsybakov (2011), Cai and Zhang (2013b,a, 2014b). However, the RIP framework is not well suited for the ROP model and would lead to sub-optimal results. See Section 2.2.2 for more discussions on the RIP and other conditions used in the literature. See also Candès et al. (2013). In this section, we introduce a Restricted Uniform Boundedness (RUB) condition which will be shown to guarantee the exact recovery of low-rank matrices in the noiseless case and stable recovery in the noisy case through the constrained nuclear norm minimization. It will also be shown that the RUB condition are satisfied by a range of random linear maps with high probability.

Definition 2.2.1 (Restricted Uniform Boundedness). *For a linear map $\mathcal{X} : \mathbb{R}^{p_1 \times p_2} \rightarrow \mathbb{R}^n$, if there exist uniform constants C_1 and C_2 such that for all nonzero rank- r matrices $A \in \mathbb{R}^{p_1 \times p_2}$*

$$C_1 \leq \frac{\|\mathcal{X}(A)\|_1/n}{\|A\|_F} \leq C_2,$$

where $\|\cdot\|_1$ means the vector ℓ_1 norm, then we say that \mathcal{X} satisfies the restricted uniform boundedness (RUB) condition of order r and constants C_1 and C_2 .

In the noiseless case, we observe $y = \mathcal{X}(A)$ and estimate the matrix A through the constrained nuclear norm minimization

$$A_* = \arg \min_M \{\|M\|_* : \mathcal{X}(M) = y\}. \quad (2.7)$$

The following theorem shows that the RUB condition guarantees the exact recovery of all rank- r matrices.

Theorem 2.2.1. *Let $k \geq 2$ be an integer. Suppose \mathcal{X} satisfies RUB of order kr with $C_2/C_1 < \sqrt{k}$, then the nuclear norm minimization method recovers all rank- r matrices. That is, for all rank- r matrices A and $y = \mathcal{X}(A)$, we have $A_* = A$, where A_* is given by (2.7).*

Theorem 2.2.1 shows that RUB of order kr with $C_2/C_1 < \sqrt{k}$ is a sufficient identifiability condition for the low-rank matrix recovery model (2.1) in the noisy case. The following result shows that the RUB condition is satisfied with high probability under the ROP model with a sufficient number of measurements.

Theorem 2.2.2. *Suppose $\mathcal{X} : \mathbb{R}^{p_1 \times p_2} \rightarrow \mathbb{R}^n$ is ROP from the standard normal distribution. For integer $k \geq 2$, positive numbers $C_1 < \frac{1}{3}$ and $C_2 > 1$, there exist constants C and δ , not depending on p_1, p_2 , and r , such that if*

$$n \geq Cr(p_1 + p_2), \quad (2.8)$$

then with probability at least $1 - e^{-n\delta}$, \mathcal{X} satisfies RUB of order kr and constants C_1 and C_2 .

Remark 2.2.1. The condition $n \geq O(r(p_1 + p_2))$ on the number of measurements is indeed necessary for \mathcal{X} to satisfy non-trivial RUB with $C_1 > 0$. Note that the degree of freedom of all rank- r matrices of $\mathbb{R}^{p_1 \times p_2}$ is $r(p_1 + p_2 - r) \geq \frac{1}{2}r(p_1 + p_2)$. If $n < \frac{1}{2}r(p_1 + p_2)$, there must exist a non-zero rank- r matrix $A \in \mathbb{R}^{p_1 \times p_2}$ such that $\mathcal{X}(A) = 0$, which leads to the failure of any non-trivial RUB for \mathcal{X} .

As a direct consequence of Theorems 2.2.1 and 2.2.2, ROP with the number of measurements $n \geq Cr(p_1 + p_2)$ guarantees the exact recovery of all rank- r matrices with high probability.

Corollary 2.2.1. *Suppose $\mathcal{X} : \mathbb{R}^{p_1 \times p_2} \rightarrow \mathbb{R}^n$ is ROP from the standard normal distribution. There exist uniform constants C and δ such that, whenever $n \geq Cr(p_1 + p_2)$, the nuclear norm minimization estimator A_* given in (2.7) recovers all rank- r matrices $A \in \mathbb{R}^{p_1 \times p_2}$ exactly with probability at least $1 - e^{-n\delta}$.*

Note that the required number of measurements $O(r(p_1 + p_2))$ above is rate optimal, since the degree of freedom for a matrix $A \in \mathbb{R}^{p_1 \times p_2}$ of rank r is $r(p_1 + p_2 - r)$, and thus at least $r(p_1 + p_2 - r)$ measurements are needed in order to recover A exactly using any method.

2.2.2 RUB, RIP and Other Conditions

We have shown that RUB implies exact recovery in the noiseless and proved that the random rank-one projections satisfy RUB with high probability whenever the number of measurements $n \geq Cr(p_1 + p_2)$. As mentioned earlier, other conditions, including the restricted isometry property (RIP), RIP in expectation, and spherical section property (SSP), have been introduced for low-rank matrix recovery based on

linear measurements. Among them, RIP is perhaps the most widely used. A linear map $\mathcal{X} : \mathbb{R}^{p_1 \times p_2} \rightarrow \mathbb{R}^n$ is said to satisfy RIP of order r with positive constants C_1 and C_2 if

$$C_1 \leq \frac{\|\mathcal{X}(A)\|_2/\sqrt{n}}{\|A\|_F} \leq C_2$$

for all rank- r matrices A . Many results have been given for low-rank matrices under the RIP framework. For example, Recht et al. (2010) showed that Gaussian ensembles satisfy RIP with high probability under certain conditions on the dimensions. Candès and Plan (2011) provided a lower bound and oracle inequality under the RIP condition. Cai and Zhang (2013b,a, 2014b) established the sharp bounds for the RIP conditions that guarantee accurate recovery of low-rank matrices.

However, the RIP framework is not suitable for the ROP model considered in the present chapter. The following lemma is proved in the Supplement.

Lemma 2.2.1. *Suppose $\mathcal{X} : \mathbb{R}^{p_1 \times p_2} \rightarrow \mathbb{R}^n$ is ROP from the standard normal distribution. Let*

$$C_1 = \min_{A:\text{rank}(A)=1} \frac{\|\mathcal{X}(A)\|_2/\sqrt{n}}{\|A\|_F} \quad \text{and} \quad C_2 = \max_{A:\text{rank}(A)=1} \frac{\|\mathcal{X}(A)\|_2/\sqrt{n}}{\|A\|_F}.$$

Then for all $t > 1$, $C_2/C_1 \geq \sqrt{p_1 p_2 / (4tn)}$ with probability at least $1 - e^{-p_1/4} - e^{-p_2/4} - \frac{8}{n(t-1)^2}$.

Lemma 2.2.1 implies that at least $O(p_1 p_2)$ number of measurements are needed in order to ensure that \mathcal{X} satisfies the RIP condition that guarantees the recovery of only rank-one matrices. Since $O(p_1 p_2)$ is the degree of freedom for all matrices $A \in \mathbb{R}^{p_1 \times p_2}$ and it is the number of measurements needed to recover all $p_1 \times p_2$ matrices (not just the low-rank matrices), Lemma 2.1 shows that the RIP framework is not suitable for the ROP model. In comparison, Theorem 2.2.2 shows that if $n \geq O(r(p_1 + p_2))$, then with high probability \mathcal{X} satisfies the RUB condition of order r with bounded C_2/C_1

, which ensures the exact recovery of all rank- r matrices.

The main technical reason for the failure of RIP under the ROP model is that RIP requires an upper bound for

$$\max_{A \in \mathcal{C}} \|\mathcal{X}(A)\|_2^2/n = \max_{A \in \mathcal{C}} \left(\sum_{j=1}^n ((\beta^{(j)})^\top A \gamma^{(j)})^2 \right) /n \quad (2.9)$$

where \mathcal{C} is a set containing low-rank matrices. The right-hand side of (2.9) involves the 4th power of the Gaussian (or sub-Gaussian) variables $\beta^{(j)}$ and $\gamma^{(j)}$. A much larger n than the bound given in (2.8) is needed in order for the linear map \mathcal{X} to satisfy the required RIP condition, which would lead to sub-optimal result.

Koltchinskii et al. (2011) uses RIP in expectation, which is a weaker condition than RIP. A random linear map $\mathcal{X} : \mathbb{R}^{p_1 \times p_2} \rightarrow \mathbb{R}^n$ is said to satisfy RIP in expectation of order r with parameters $0 < \mu < \infty$ and $0 \leq \delta_r < 1$ if

$$(1 - \delta_r) \|A\|_F^2 \leq \mu \frac{1}{n} E \|\mathcal{X}(A)\|_2^2 \leq (1 + \delta_r) \|A\|_F^2$$

for all rank- r matrices $A \in \mathbb{R}^{p_1 \times p_2}$. This condition was originally introduced by Koltchinskii et al. (2011) to prove an oracle inequality for the estimator they proposed and a minimax lower bound. The condition is not sufficiently strong to guarantee the exact recovery of rank- r matrices in the noiseless case. To be more specific, the bounds in Theorems 1 and 2 in Koltchinskii et al. (2011) depend on $\mathbf{M} = \left\| \frac{1}{n} \sum_{i=1}^n (y_i X_i - E(y_i X_i)) \right\|$, which might be non-zero even in the noiseless case. In fact, in the ROP model considered in the present chapter, we have

$$\begin{aligned} \frac{1}{n} E \|\mathcal{X}\|_2^2 &= \frac{1}{n} \sum_{i=1}^n E (\beta^{(i)T} A \gamma^{(i)})^2 = E(\beta^\top A \gamma \gamma^\top A^\top \beta) \\ &= E \text{tr}(A \gamma \gamma^\top A^\top \beta \beta^\top) = \text{tr}(A A^\top) = \|A\|_F^2 \end{aligned}$$

which means RIP in expectation is met for $\mu = 1$ and $\delta_r = 0$ for any number of measurements n . However, as we discussed earlier in this section that at least $O(r(p_1 + p_2))$ measurements are needed to guarantee the model identifiability for recovery of all rank- r matrices, we can see that RIP in expectation cannot ensure recovery.

Dvijotham and Fazel (2010) and Oymak et al. (2011) used a condition called the spherical section property (SSP) which focuses on the null space of \mathcal{X} . $\text{Null}(\mathcal{X})$ is said to satisfy Δ -SSP if for all $Z \in \text{Null}(\mathcal{X}) \setminus \{0\}$, $\|Z\|_* / \|Z\|_F \geq \sqrt{\Delta}$. Dvijotham and Fazel (2010) showed that if \mathcal{X} satisfies Δ -SSP, $p_1 \leq p_2$ and $\text{rank}(A) < \min\left(3p_1/4 - \sqrt{9p_1^2/16 - p_1\Delta/4}, p_1/2\right)$, the nuclear norm minimization (2.7) recovers A exactly in the noiseless case. However, the SSP condition is difficult to utilize in the ROP framework since it is hard to characterize the matrices $Z \in \text{Null}(\mathcal{X})$ when \mathcal{X} is rank-one projections.

2.2.3 Gaussian Noise Case

We now turn to the Gaussian noise case where $z_i \stackrel{iid}{\sim} N(0, \sigma^2)$ in (2.5). We begin by introducing a constrained nuclear norm minimization estimator. Define two sets

$$\mathcal{Z}_1 = \{z : \|z\|_1/n \leq \sigma\} \quad \text{and} \quad \mathcal{Z}_2 = \{z : \|\mathcal{X}^*(z)\| \leq \eta\} \quad (2.10)$$

where $\eta = \sigma \left(12\sqrt{\log n}(p_1 + p_2) + 6\sqrt{2n(p_1 + p_2)}\right)$, and let

$$\mathcal{Z}_G = \mathcal{Z}_1 \cap \mathcal{Z}_2. \quad (2.11)$$

Note that both \mathcal{Z}_1 and \mathcal{Z}_2 are convex sets and so is \mathcal{Z}_G . Our estimator of A is given by

$$\hat{A} = \arg \min_M \{\|M\|_* : y - \mathcal{X}(M) \in \mathcal{Z}_G\}. \quad (2.12)$$

The following theorem gives the rate of convergence for the estimator \hat{A} under the squared Frobenius norm loss.

Theorem 2.2.3 (Upper Bound). *Let \mathcal{X} be ROP from the standard normal distribution and let $z_1, \dots, z_n \stackrel{iid}{\sim} N(0, \sigma^2)$. Then there exist uniform constants C, W and δ such that, whenever $n \geq Cr(p_1 + p_2)$, the estimator \hat{A} given in (2.12) satisfies*

$$\|\hat{A} - A\|_F^2 \leq W\sigma^2 \min\left(\frac{r \log n(p_1 + p_2)^2}{n^2} + \frac{r(p_1 + p_2)}{n}, 1\right) \quad (2.13)$$

for all rank- r matrices A , with probability at least $1 - 11/n - 3 \exp(-\delta(p_1 + p_2))$.

Moreover, we have the following lower bound result for ROP.

Theorem 2.2.4 (Lower Bound). *Assume that \mathcal{X} is ROP from the standard normal distribution and that $z_1, \dots, z_n \stackrel{iid}{\sim} N(0, \sigma^2)$. There exists a uniform constant C such that, when $n > Cr \max(p_1, p_2)$, with probability at least $1 - 26n^{-1}$,*

$$\inf_{\hat{A}} \sup_{A \in \mathbb{R}^{p_1 \times p_2}: \text{rank}(A)=r} P_z \left(\|\hat{A} - A\|_F^2 \geq \frac{\sigma^2 r(p_1 + p_2)}{32n} \right) \geq 1 - e^{-(p_1 + p_2)r/64} \quad (2.14)$$

$$\inf_{\hat{A}} \sup_{A \in \mathbb{R}^{p_1 \times p_2}: \text{rank}(A)=r} E_z \|\hat{A} - A\|_F^2 \geq \frac{\sigma^2 r(p_1 + p_2)}{4n} \quad (2.15)$$

where E_z , and P_z are the expectation and probability with respect to the distribution of z .

When $n < r \max(p_1, p_2)$, then

$$\inf_{\hat{A}} \sup_{A \in \mathbb{R}^{p_1 \times p_2}: \text{rank}(A)=r} E_z \|\hat{A} - A\|_F^2 = \infty. \quad (2.16)$$

Comparing Theorem 2.2.3 and Theorem 3.3.3, our proposed estimator is rate optimal in the Gaussian noise case when $n \gtrsim \log n(p_1 + p_2)$ (which is equivalent to $n \gtrsim (p_1 + p_2) \log(p_1 + p_2)$) or $n \sim r(p_1 + p_2)$. Since $n \gtrsim r(p_1 + p_2)$, this condition

is also implied by $r \gtrsim \log(p_1 + p_2)$. Theorem 3.3.3 also shows that no method can recover matrices of rank r consistently if the number of measurements n is smaller than $r \max(p_1, p_2)$.

The result in Theorem 2.2.3 can also be extended to the more general case where the matrix of interest A is only approximately low-rank. Let $A = A_{\max(r)} + A_{-\max(r)}$.

Proposition 2.2.1. *Under the assumptions of Theorem 2.2.3, there exist uniform constants C, W_1, W_2 and δ such that, whenever $n \geq Cr(p_1 + p_2)$, the estimator \hat{A} given in (2.12) satisfies*

$$\|\hat{A} - A\|_F^2 \leq W_1 \sigma^2 \min\left(\frac{r \log n (p_1 + p_2)^2}{n^2} + \frac{r(p_1 + p_2)}{n}, 1\right) + W_2 \frac{\|A_{-\max(r)}\|_*^2}{r} \quad (2.17)$$

for all matrices $A \in \mathbb{R}^{p_1 \times p_2}$, with probability at least $1 - 11/n - 3 \exp(-\delta(p_1 + p_2))$.

If the matrix A is approximately of rank r , then $\|A_{-\max(r)}\|_*$ is small, and the estimator \hat{A} continues to perform well. This result shows that the constrained nuclear norm minimization estimator is adaptive to the rank r and robust against perturbations of small amplitude.

Remark 2.2.2. All the results remain true if the Gaussian design is replaced by the Rademacher design where entries of $\beta^{(i)}$ and $\gamma^{(i)}$ are i.i.d. ± 1 with probability $\frac{1}{2}$. More general sub-Gaussian design case will be discussed in Section 2.3.

Remark 2.2.3. The estimator \hat{A} we propose here is the minimizer of the nuclear norm under the constraint of the intersection of two convex sets \mathcal{Z}_1 and \mathcal{Z}_2 . Nuclear norm minimization under either one of the two constraints, called “ ℓ_1 constraint nuclear norm minimization” ($\mathcal{Z} = \mathcal{Z}_1$) and “matrix Dantzig Selector” ($\mathcal{Z} = \mathcal{Z}_2$), has been studied before in various settings (Candès and Plan 2011, Recht et al. 2010, Cai and Zhang 2013b,a, 2014b, Chen et al. 2013). Our analysis indicates the following.

1. The ℓ_1 constraint minimization performs better than the matrix Dantzig Selector for small n ($n \sim r(p_1 + p_2)$) when $r \ll \log n$;
2. The matrix Dantzig Selector outperforms the ℓ_1 constraint minimization for large n as the loss of the matrix Dantzig Selector decays at the rate $O(n^{-1})$;
3. The proposed estimator \hat{A} combines the advantages of the two estimators.

See Section 2.5 for a comparison of numerical performances of the three methods.

2.2.4 Recovery of Symmetric Matrices

For applications such as low-dimensional Euclidean embedding (Trosset 2000, Recht et al. 2010), phase retrieval (Candès et al. 2013, Candès et al. 2011), and covariance matrix estimation (Chen et al. 2013, Cai et al. 2013b,a), the low-rank matrix A of interest is known to be symmetric. Examples of such matrices include distance matrices, Gram matrices, and covariance matrices. When the matrix A is known to be symmetric, the ROP design can be further simplified by taking $\beta^{(i)} = \gamma^{(i)}$.

Denote by \mathbb{S}^p the set of all $p \times p$ symmetric matrices in $\mathbb{R}^{p \times p}$. Let $\beta^{(1)}, \beta^{(2)}, \dots, \beta^{(n)}$ be independent p -dimensional random vectors with i.i.d. entries generated from some distribution \mathcal{P} . Define a linear map $\mathcal{X} : \mathbb{S}^p \rightarrow \mathbb{R}^n$ by

$$[\mathcal{X}(A)]_i = (\beta^{(i)})^\top A \beta^{(i)}, \quad i = 1, \dots, n.$$

We call such a linear map \mathcal{X} “Symmetric Rank-One Projections” (SROP) from the distribution \mathcal{P} .

Suppose we observe

$$y_i = (\beta^{(i)})^\top A \beta^{(i)} + z_i, \quad i = 1, \dots, n \tag{2.18}$$

and wish to recover the symmetric matrix A . As for the ROP model, in the noiseless case we estimate A under the SROP model by

$$A_* = \arg \min_{M \in \mathbb{S}^p} \{\|M\|_* : y = \mathcal{X}(M)\}. \quad (2.19)$$

Proposition 2.2.2. *Let \mathcal{X} be SROP from the standard normal distribution. Similar to Corollary 2.2.1, there exist uniform constants C and δ such that, whenever $n \geq Crp$, the nuclear norm minimization estimator A_* given by (2.19) recovers exactly all rank- r symmetric matrices $A \in \mathbb{S}^p$ with probability at least $1 - e^{-n\delta}$.*

For the noisy case, we propose a constraint nuclear norm minimization estimator similar to (2.12). Define the linear map $\tilde{\mathcal{X}} : \mathbb{R}^{p_1 \times p_2} \rightarrow \mathbb{R}^{\lfloor \frac{n}{2} \rfloor}$ by

$$[\tilde{\mathcal{X}}(A)]_i = [\mathcal{X}(A)]_{2i-1} - [\mathcal{X}(A)]_{2i}, \quad i = 1, \dots, \lfloor \frac{n}{2} \rfloor \quad (2.20)$$

and define $\tilde{y} \in \mathbb{R}^{\lfloor n/2 \rfloor}$ by

$$\tilde{y}_i = y_{2i-1} - y_{2i}, \quad i = 1, \dots, \lfloor \frac{n}{2} \rfloor. \quad (2.21)$$

Based on the definition of $\tilde{\mathcal{X}}$, the dual map $\tilde{\mathcal{X}}^* : \mathbb{R}^{\lfloor \frac{n}{2} \rfloor} \rightarrow \mathbb{S}^p$ is

$$\tilde{\mathcal{X}}^*(z) = \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} z_i (\beta^{(2i-1)} \beta^{(2i-1)\top} - \beta^{(2i)} \beta^{(2i)\top}) \quad (2.22)$$

Let $\eta = 24\sigma (\sqrt{pn} + 2p\sqrt{2 \log n})$. The estimator \hat{A} of the matrix A is given by

$$\hat{A} = \arg \min_{M \in \mathbb{S}^p} \left\{ \|M\|_* : \|y - \mathcal{X}(M)\|_1/n \leq \sigma, \|\tilde{\mathcal{X}}^*(\tilde{y} - \tilde{\mathcal{X}}(M))\| \leq \eta \right\}. \quad (2.23)$$

Remark 2.2.4. An important property in the ROP model considered in Section 2.2.3 is that $E\mathcal{X} = 0$, i.e., $EX_i = 0$ for all the measurement matrices X_i . However, under

the SROP model $X_i = \beta^{(i)}(\beta^{(i)})^\top$ and so $E\mathcal{X} \neq 0$. The step of taking the pairwise differences in (2.20) and (2.21) is to ensure that $E\tilde{\mathcal{X}} = 0$.

The following result is similar to the upper bound given in Proposition 2.2.1 for ROP.

Proposition 2.2.3. *Let \mathcal{X} be SROP from the standard normal distribution and let $z_1, \dots, z_n \stackrel{iid}{\sim} N(0, \sigma^2)$. There exist constants C, W_1, W_2 and δ such that, whenever $n \geq Crp$, the estimator \hat{A} given in (2.23) satisfies*

$$\|\hat{A} - A\|_F^2 \leq W_1 \sigma^2 \min\left(\frac{rp^2 \log n}{n^2} + \frac{rp}{n}, 1\right) + W_2 \frac{\|A_{-\max(r)}\|_*^2}{r} \quad (2.24)$$

for all matrices $A \in \mathbb{S}^p$, with probability at least $1 - 15/n - 5 \exp(-p\delta)$.

In addition, we also have lower bounds for SROP, which show that the proposed estimator is rate-optimal when $n \gtrsim p \log n$ or $n \sim rp$, and no estimator can recover a rank- r matrix consistently if the number of measurements $n < \lfloor \frac{r}{2} \rfloor \cdot \lfloor \frac{p}{2} \rfloor$.

Proposition 2.2.4 (Lower Bound). *Assume that \mathcal{X} is SROP from the standard normal distribution and that $z_1, \dots, z_n \stackrel{iid}{\sim} N(0, \sigma^2)$. Then there exists a uniform constant C such that, when $n > Crp$ and $p, r \geq 2$, with probability at least $1 - 26n^{-1}$,*

$$\inf_{\hat{A}} \sup_{A \in \mathbb{S}^p: \text{rank}(A)=r} P_z \left(\|\hat{A} - A\|_F^2 \geq \frac{\sigma^2 rp}{192n} \right) \geq 1 - e^{-pr/192}$$

$$\inf_{\hat{A}} \sup_{A \in \mathbb{S}^p: \text{rank}(A)=r} E_z \|\hat{A} - A\|_F^2 \geq \frac{\sigma^2 rp}{24n}$$

where \hat{A} is any estimator of A , E_z, P_z are the expectation and probability with respect to z .

When $n < \lfloor \frac{r}{2} \rfloor \cdot \lfloor \frac{p}{2} \rfloor$ and $p, r \geq 2$, then

$$\inf_{\hat{A}} \sup_{A \in \mathbb{S}^p: \text{rank}(A)=r} E_z \|\hat{A} - A\|_F^2 = \infty.$$

2.3 Sub-Gaussian Design and Sub-Gaussian Noise

We have focused on the Gaussian design and Gaussian noise distribution in Section 2.2. These results can be further extended to more general distributions. In this section we consider the case where the ROP design is from a symmetric sub-Gaussian distribution \mathcal{P} and the errors z_i are also from a sub-Gaussian distribution. We say the distribution of a random variable Z is sub-Gaussian with parameter τ if

$$P(|Z| \geq t) \leq 2 \exp(-t^2/(2\tau^2)), \quad \text{for all } t > 0. \quad (2.25)$$

The following lemma provides a necessary and sufficient condition for symmetric sub-Gaussian distributions.

Lemma 2.3.1. *Let \mathcal{P} be a symmetric distribution and let the random variable $X \sim \mathcal{P}$.*

Define

$$\alpha_{\mathcal{P}} = \sup_{k \geq 1} \left(\frac{EX^{2k}}{(2k-1)!!} \right)^{\frac{1}{2k}}. \quad (2.26)$$

Then the distribution \mathcal{P} is sub-Gaussian if and only if $\alpha_{\mathcal{P}}$ is finite.

For the sub-Gaussian ROP design and sub-Gaussian noise, we estimate the low-rank matrix A by the estimator \hat{A} given in (3.3) with

$$\begin{aligned} \mathcal{Z}_G = & \{z : \|z\|_1/n \leq 6\tau\} \\ & \cap \left\{ z : \|\mathcal{X}^*(z)\| \leq 6\alpha_{\mathcal{P}}^2\tau \left(\sqrt{6n(p_1 + p_2)} + 2\sqrt{\log n(p_1 + p_2)} \right) \right\} \end{aligned} \quad (2.27)$$

where $\alpha_{\mathcal{P}}$ is given in (2.26).

Theorem 2.3.1. *Suppose $\mathcal{X} : \mathbb{R}^{p_1 \times p_2} \rightarrow \mathbb{R}^n$ is ROP from a symmetric and variance 1 sub-Gaussian distribution \mathcal{P} . Assume that z_i are i.i.d. sub-Gaussian with parameter τ and \hat{A} is given by (3.3) with $\mathcal{Z} = \mathcal{Z}_G$ defined in (2.27). Then there exist constants C, W_1, W_2, δ which only depend on \mathcal{P} , such that if $n \geq Cr(p_1 + p_2)$, we have*

$$\|\hat{A} - A\|_F^2 \leq W_1 \tau^2 \min \left(\frac{r \log n (p_1 + p_2)^2}{n^2} + \frac{r(p_1 + p_2)}{n}, 1 \right) + W_2 \frac{\|A_{-\max(r)}\|_*^2}{r} \quad (2.28)$$

with probability at least $1 - 2/n - 5e^{-\delta(p_1 + p_2)}$.

An exact recovery result in the noiseless case for the sub-Gaussian design follows directly from Theorem 2.3.1. If $z = 0$, then, with high probability, all rank- r matrices A can be recovered exactly via the constrained nuclear minimization (2.7) whenever $n \geq C_{\mathcal{P}} r(p_1 + p_2)$ for some constant $C_{\mathcal{P}} > 0$.

Remark 2.3.1. For the SROP model considered in Section 2.2.4, we can similarly extend the results to the case of sub-Gaussian design and sub-Gaussian noise. Suppose \mathcal{X} is SROP from a symmetric variance 1 sub-Gaussian distribution \mathcal{P} (other than the Rademacher ± 1 distribution) and z satisfies (2.25). Define the estimator of the low-rank matrix A by

$$\hat{A} = \arg \min_{M \in \mathbb{S}^p} \left\{ \|M\|_* : \|y - \mathcal{X}(M)\|_1/n \leq 6\tau, \|\tilde{\mathcal{X}}^*(\tilde{y} - \tilde{\mathcal{X}}(M))\| \leq \eta \right\} \quad (2.29)$$

where $\eta = C_{\mathcal{P}} (\sqrt{np} + \sqrt{\log np})$ with $C_{\mathcal{P}}$ some constant depending on \mathcal{P} .

Proposition 2.3.1. *Suppose $\mathcal{X} : \mathbb{R}^{p \times p} \rightarrow \mathbb{R}^n$ is SROP from a symmetric sub-Gaussian distribution \mathcal{P} with variance 1. Also, assume that $\text{Var}(\mathcal{P}^2) > 0$ (i.e. $\text{Var}(w^2) > 0$ where $w \sim \mathcal{P}$). Let \hat{A} be given by (2.29). Then there exist constants*

$C, C_{\mathcal{P}}, W_1, W_2$, and δ which only depend on \mathcal{P} , such that for $n \geq Crp$,

$$\|\hat{A} - A\|_F^2 \leq W_1 \tau^2 \min\left(\frac{rp^2 \log n}{n^2} + \frac{rp}{n}, 1\right) + W_2 \frac{\|A_{-\max(r)}\|_*^2}{r} \quad (2.30)$$

with probability at least $1 - 2/n - 5e^{-\delta p}$.

By restricting $\text{Var}(\mathcal{P}^2) > 0$, Rademacher ± 1 is the only symmetric and variance 1 distribution that has been excluded. The reason why the Rademacher ± 1 distribution is an exception for the SROP design is as follows. If $\beta^{(i)}$ are i.i.d. Rademacher ± 1 distributed, then

$$[\mathcal{X}(A)]_i = (\beta^{(i)})^\top A \beta^{(i)} = \sum_{j=1}^p a_{jj} + \sum_{j \neq k} \beta_j^{(i)} \beta_k^{(i)} a_{jk}, \quad i = 1, \dots, n.$$

So the only information contained in $\mathcal{X}(A)$ about $\text{diag}(A)$ is $\text{trace}(A)$, which makes it impossible to recover the whole matrix A .

2.4 Application to Estimation of Spiked Covariance Matrix

In this section, we consider an interesting application of the methods and results developed in the previous sections to estimation of a spiked covariance matrix based on one-dimensional projections. As mentioned in the introduction, spiked covariance matrix model has been used in a wide range of applications and it has been well studied in the context of PCA based on i.i.d. data where one observes i.i.d. p -dimensional random vectors $X^{(1)}, \dots, X^{(n)}$ with mean 0 and covariance matrix Σ , where $\Sigma = I_p + \Sigma_0$ and Σ_0 being low-rank. See, for example, (Johnstone 2001, Birnbaum et al. 2013, Cai et al. 2013b,a). Here we consider estimation of Σ_0 (or equivalently Σ) based only on one-dimensional random projections of $X^{(i)}$. More specifically, suppose that the

random vectors $X^{(1)}, \dots, X^{(n)}$ are not directly observable and instead we observe

$$\xi_i = \langle \beta^{(i)}, X^{(i)} \rangle = \sum_{j=1}^p \beta_j^{(i)} X_j^{(i)}, \quad i = 1, \dots, n, \quad (2.31)$$

where $\beta^{(i)} \stackrel{iid}{\sim} N(0, I_p)$. The goal is to recover Σ_0 from the projections $\{\xi_i, i = 1, \dots, n\}$.

Let $y = (y_1, \dots, y_n)^\top$ with $y_i = \xi_i^2 - \beta^{(i)\top} \beta^{(i)}$. Note that

$$E(\xi^2 | \beta) = E\left(\sum_{i,j} \beta_i \beta_j X_i X_j | \beta\right) = \sum_{i,j} \beta_i \beta_j \sigma_{i,j} = \beta^\top \Sigma \beta$$

and so $E(\xi^2 - \beta^\top \Sigma \beta | \beta) = \beta^\top \Sigma_0 \beta$. Define a linear map $\mathcal{X} : \mathbb{S}^p \rightarrow \mathbb{R}^n$ by

$$[\mathcal{X}(A)]_i = \beta^{(i)\top} A \beta^{(i)}. \quad (2.32)$$

Then y can be formally written as

$$y = \mathcal{X}(\Sigma_0) + z \quad (2.33)$$

where $z = y - \mathcal{X}(\Sigma_0)$. We define the corresponding $\tilde{\mathcal{X}}$ and \tilde{y} as in (2.20) and (2.21) respectively, and apply the constraint nuclear norm minimization to recover the low-rank matrix Σ_0 by

$$\hat{\Sigma}_0 = \arg \min_M \left\{ \|M\|_* : \|y - \mathcal{X}(M)\| \leq \eta_1, \|\tilde{\mathcal{X}}^*(\tilde{y} - \tilde{\mathcal{X}}(M))\| \leq \eta_2 \right\}. \quad (2.34)$$

The tuning parameters η_1 and η_2 are chosen as

$$\eta_1 = c_1 \sum_{i=1}^n \xi_i^2 \quad \text{and} \quad \eta_2 = 24c_2 \sqrt{p \sum_{i=1}^n \xi_i^4 + 48c_3 p \log n \max_{1 \leq i \leq n} \xi_i^2} \quad (2.35)$$

where $c_1 > \sqrt{2}$, $c_2, c_3 > 1$ are constants.

We have the following result on the estimator (2.34) for spiked covariance matrix estimation.

Theorem 2.4.1. *Suppose $n \geq 3$, we observe $\xi_i, i = 1, \dots, n$, as in (2.31), where $\beta^{(i)} \stackrel{iid}{\sim} N(0, I_p)$ and $X^{(1)}, \dots, X^{(n)} \stackrel{iid}{\sim} N(0, \Sigma)$ with $\Sigma = I_p + \Sigma_0$ and Σ_0 positive semidefinite and $\text{rank}(\Sigma_0) \leq r$. Let $\hat{\Sigma}_0$ be given by (2.34). Then there exist uniform constants C, D, δ such that when $n \geq Drp$,*

$$\|\hat{\Sigma}_0 - \Sigma_0\|_F^2 \leq C \min \left(\frac{rp}{n} \|\Sigma\|_*^2 + \frac{rp^2 \log^4 n}{n^2} (\|\Sigma\|_*^2 + \log^2 n \|\Sigma\|^2), \|\Sigma\|_*^2 \right) \quad (2.36)$$

with probability at least $1 - O(1/n) - 4 \exp(-p\delta) - \frac{2}{\sqrt{2\pi \log n}}$.

Remark 2.4.1. We have focused estimation of spiked covariance matrices on the setting where the random vectors $X^{(i)}$ are Gaussian. Similar to the discussion in Section 2.3, the results given here can be extended to more general distributions under certain moment conditions.

Remark 2.4.2. The problem considered in this section is related to the so-called covariance sketching problem considered in Dasarthy et al. (2012). In covariance sketching, the goal is to estimate the covariance matrix of high-dimensional random vectors $X^{(1)}, \dots, X^{(n)}$ based on the low dimensional projections

$$y^{(i)} = QX^{(i)}, \quad i = 1, \dots, n,$$

where Q is a fixed $m \times p$ projection matrix with $m < p$. The main differences between the two settings are that the projection matrix in covariance sketch is the same for all $X^{(i)}$ and the dimension m is still relatively large with $m \geq C\sqrt{p} \log^3 p$ for some $C > 0$. In our setting, $m = 1$ and Q is random and varies with i . The techniques for

solving the two problems are very different. Comparing to Dasarathy et al. (2012), the results in this section indicate that there is a significant advantage to have different random projections for different random vectors $X^{(i)}$ as opposed to having the same projection for all $X^{(i)}$.

2.5 Simulation Results

The constrained nuclear norm minimization methods can be efficiently implemented. The estimator \hat{A} proposed in Section 2.2.3 can be implemented by the following convex programming:

$$\begin{aligned} & \text{minimize} && \text{Tr}(B_1) + \text{Tr}(B_2) \\ & \text{subject to} && \begin{bmatrix} B_1 & A \\ A^\top & B_2 \end{bmatrix} \succeq 0, \quad \|y - \mathcal{X}(A)\|_1 \leq \lambda_1, \quad \|\mathcal{X}^*(y - \mathcal{X}(A))\| \leq \lambda_2, \end{aligned} \quad (2.37)$$

with optimization variables $B_1 \in \mathbb{S}^{p_1}$, $B_2 \in \mathbb{S}^{p_2}$, $A \in \mathbb{R}^{p_1 \times p_2}$. We use the CVX package (Grant and Boyd, 2012, 2008) to implement the proposed procedures. In this section, a simulation study is carried out to investigate the numerical performance of the proposed procedures for low-rank matrix recovery in various settings.

We begin with the noiseless case. In this setting, Theorem 2.2.2 and Corollary 2.2.1 show that the nuclear norm minimization recovers a rank r matrix exactly whenever

$$n \geq Cr \max(p_1, p_2). \quad (2.38)$$

A similar result holds for the Gaussian ensemble (Candès and Plan, 2011). However, the minimum constant C that guarantees the exact recovery with high probability is not specified in either case. It is of practical interest to find the minimum constant C . For this purpose, we randomly generate $p_1 \times p_2$ rank r matrices A as $A =$

$X^T Y$, where $X \in \mathbb{R}^{r \times p_1}$, $Y \in \mathbb{R}^{r \times p_2}$ are i.i.d. Gaussian matrices. We compare ROP from the standard Gaussian distribution and the Gaussian ensemble, with the number of measurements $n = Cr \max(p_1, p_2)$ from a range of values of C using the constrained nuclear norm minimization (2.7). A recovery is considered successful if $\|\hat{A} - A\|_F / \|A\|_F \leq 10^{-4}$. Figure 2.1 shows the rate of successful recovery when $p_1 = p_2 = 100$ and $r = 5$.

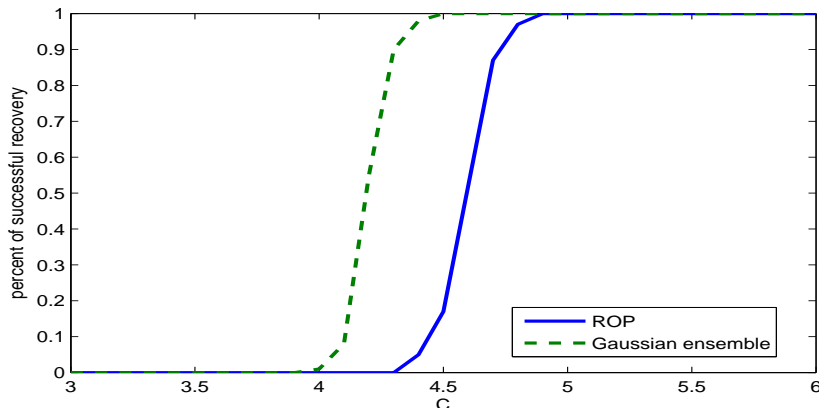


Figure 2.1: Rates of successful recovery for the ROP and Gaussian ensemble with $p_1 = p_2 = 100$, $r = 5$, and $n = Cr \max(p_1, p_2)$ for C ranging from 3 to 6.

The numerical results show that for ROP from the Gaussian distribution, the minimum constant C to ensure exact recovery with high probability is slightly less than 5 in the small scale problems ($p_1, p_2 \leq 100$) we tested. The corresponding minimum constant C for the Gaussian ensemble is about 4.5. Matrix completion requires much larger number of measurements. Based on the theoretical analyses given in Candès and Recht (2009), Recht (2011), the required number of measurements for matrix completion is $O(\mu r(p_1 + p_2) \log^2(p_1 + p_2))$, where $\mu \geq 1$ is some coherence constant describing the “spikedness” of the matrix A . Hence for matrix completion, the factor C in (2.38) needs to grow with the dimensions p_1 and p_2 and it requires $C \gtrsim \mu \log^2(p_1 + p_2)$, which is much larger than what is needed for the ROP or Gaus-

sian ensemble. The required storage space for the Gaussian ensemble is much greater than that for the ROP. In order to ensure accurate recovery of $p \times p$ matrices of rank r , one needs at least $4.5rp^3$ bytes of space to store the measurement matrices, which could be prohibitively large for the recovery of high-dimensional matrices. In contrast, the storage space for the projection vectors in ROP is only $10rp^2$ bytes, which is far smaller than what is required by the Gaussian ensemble in the high-dimensional case.

We then consider the recovery of approximately low-rank matrices to investigate the robustness of the method against small perturbations. To this end, we randomly draw 100×100 matrix A as $A = U \cdot \text{diag}(1, 2^{-1/2}, \dots, r^{-1/2}) \cdot V^\top$, where $U \in \mathbb{R}^{100 \times r}$ and $V \in \mathbb{R}^{100 \times r}$ are random matrices with orthonormal columns. We then observe $n = 2000$ random rank-one projections with the measurement vectors being i.i.d. Gaussian. Based on the observations, the nuclear minimization procedure (2.7) is applied to estimate A . The results for different values of r are shown in Figure 2.2. It can be seen from the plot that in this setting one can exactly recover a matrix of rank at most 4 with 2000 measurements. However, when the rank r of the true matrix A exceeds 4, the estimate is still stable. The theoretical result in Proposition 2.2.1 bounds the loss (solid line) at $O(\|A_{-\max(4)}\|_*^2/4)$ (shown in the dashed line) with high probability, which corresponds to Figure 2.2.

We now turn to the noisy case. The low-rank matrices A are generated by $A = X^\top Y$, where $X \in \mathbb{R}^{r \times p_1}$ and $Y \in \mathbb{R}^{r \times p_2}$ are i.i.d. Gaussian matrices. The ROP \mathcal{X} is from the standard Gaussian distribution and the noise vector $z \sim N_n(0, \sigma^2)$. Based on (\mathcal{X}, y) with $y = \mathcal{X}(A) + z$, we compare our proposed estimator \hat{A} with the ℓ_1 constraint minimization estimator \hat{A}^{ℓ_1} (Chen et al., 2013) and the matrix Dantzig

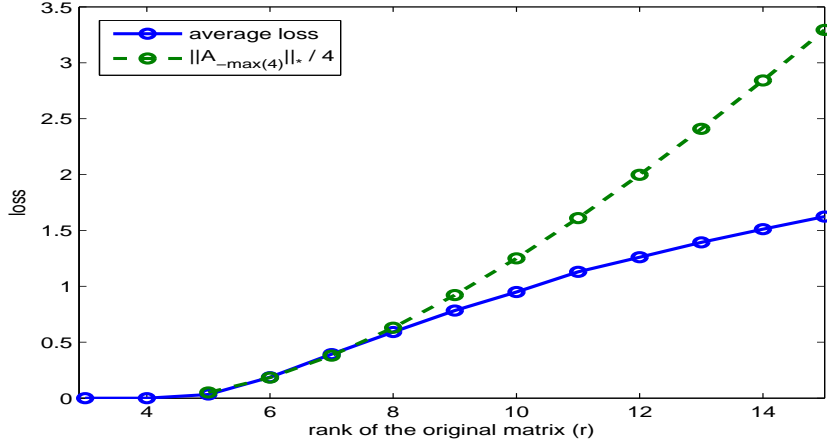


Figure 2.2: Recovery accuracy (solid line) for approximately low-rank matrices with different values of r , where $p_1 = p_2 = 100$, $n = 2000$, $\sigma(A) = (1, 1/\sqrt{2}, \dots, 1/\sqrt{r})$. The dashed line is the theoretical upper bound.

Selector \hat{A}^{DS} (Candès and Plan, 2011), where

$$\begin{aligned} \hat{A} &= \arg \min_M \{\|M\|_* : y - \mathcal{X}(M) \in \mathcal{Z}_1 \cap \mathcal{Z}_2\}, \\ \hat{A}^{\ell_1} &= \arg \min_M \{\|M\|_* : y - \mathcal{X}(M) \in \mathcal{Z}_1\}, \\ \hat{A}^{DS} &= \arg \min_M \{\|M\|_* : y - \mathcal{X}(M) \in \mathcal{Z}_2\}, \end{aligned}$$

with $\mathcal{Z}_1 = \{z : \|z\|_1/n \leq \sigma\}$ and $\mathcal{Z}_2 = \{z : \|\mathcal{X}(z)\| \leq \sigma(\sqrt{\log n}(p_1 + p_2) + \sqrt{n(p_1 + p_2)})\}$. Note that \hat{A}^{ℓ_1} is similar to the estimator proposed in Chen et al. (2013), except their estimator is for symmetric matrices under the SROP but ours is for general low-rank matrices under the ROP. Figure 2.3 compares the performance of the three estimators. It can be seen from the left panel that for small n , ℓ_1 constrained minimization outperforms the matrix Dantzig Selector, while our estimator outperforms both \hat{A}^{ℓ_1} and \hat{A}^{DS} . When n is large, our estimator and \hat{A}^{DS} are essentially the same and both outperforms \hat{A}^{ℓ_1} . The right panel of Figure 2.3 plots the

ratio of the squared Frobenius norm loss of \hat{A}^{ℓ_1} to that of our estimator. The ratio increases with n . These numerical results are consistent with the observations made in Remark 2.2.3.

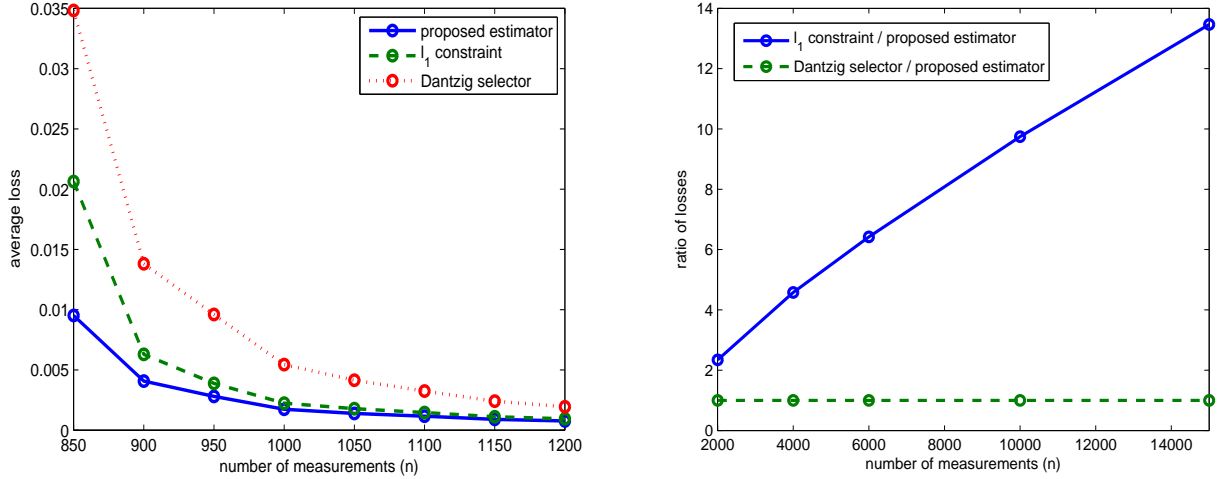


Figure 2.3: Left Panel: Comparison of the proposed estimator with \hat{A}^{ℓ_1} and \hat{A}^{DS} for $p_1 = p_2 = 50$, $r = 4$, $\sigma = 0.01$, and n ranging from 850 to 1200. Right Panel: Ratio of the squared Frobenius norm loss of \hat{A}^{ℓ_1} to that of the proposed estimator for $p_1 = p_2 = 50$, $r = 4$, and n varying from 2000 to 15000.

We now turn to the recovery of symmetric low-rank matrices under the SROP model (2.18). Let \mathcal{X} be SROP from the standard normal distribution. We consider the setting where $p = 40$, n varies from 50 to 600, $z_i \sim \sigma \cdot \mathcal{U}[-1, 1]$ with $\sigma = .1, .01, .001$ or $.0001$, and A is randomly generated as rank-5 matrix by the same procedure discussed above. The setting is identical to the one considered in Section 5.1 of Chen et al. (2013). Although we cannot exactly repeat the simulation study in Chen et al. (2013) as they did not specify the choice of the tuning parameter, we can implement both

our procedure

$$\hat{A} = \arg \min_M \left\{ \|M\|_* : \|y - \mathcal{X}(M)\|_1 \leq \frac{n\sigma}{2}, \right. \\ \left. \|\tilde{\mathcal{X}}^*(\tilde{y} - \tilde{\mathcal{X}}(M))\| \leq \frac{\sigma(\sqrt{\log np} + \sqrt{np})}{3} \right\}$$

and the estimator \hat{A}^{ℓ_1} with only the ℓ_1 constraint which was proposed by Chen et al. (2013)

$$\hat{A}^{\ell_1} = \arg \min_M \left\{ \|M\|_* : \|y - \mathcal{X}(M)\|_1 \leq \frac{n\sigma}{2} \right\}.$$

The results are given in Figure 2.4. It can be seen that our estimator \hat{A} outperforms the estimator \hat{A}^{ℓ_1} .

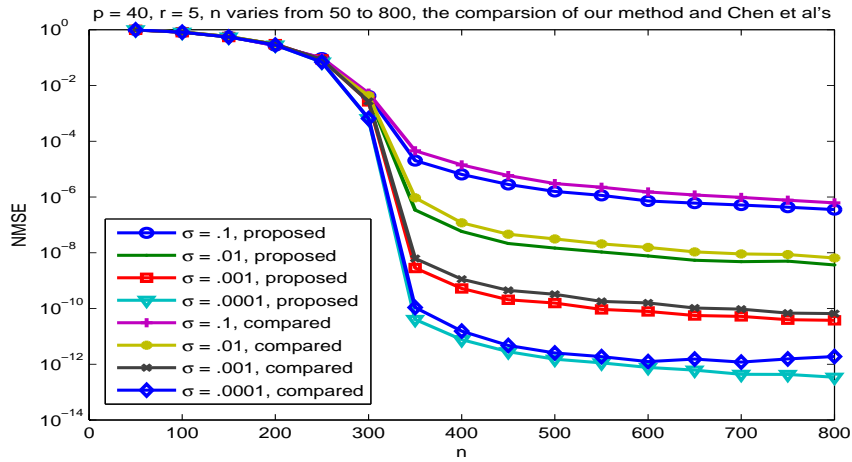


Figure 2.4: Comparison of the proposed estimator \hat{A} with the \hat{A}^{ℓ_1} . Here $p = 40$, $r = 5$, $\sigma = 0.1, 0.01, 0.001, 0.0001$, and n ranges from 50 to 800.

2.5.1 Data Driven Selection of Tuning Parameters

We have so far considered the estimators

$$\hat{A} = \arg \min_B \{ \|B\|_* : \|y - \mathcal{X}(B)\|_1/n \leq \lambda, \|\mathcal{X}^*(y - \mathcal{X}(B))\| \leq \eta \}, \quad (2.39)$$

$$\hat{A} = \arg \min_M \{ \|M\|_* : \|y - \mathcal{X}(M)\|_1/n \leq \lambda, \|\tilde{\mathcal{X}}^*(\tilde{y} - \tilde{\mathcal{X}}(M))\| \leq \eta \} \quad (2.40)$$

for the ROP and SROP, respectively. The theoretical choice of the tuning parameters λ and η depends on the knowledge of the error distribution such as the variance. In real applications, such information may not be available and/or the theoretical choice may not be the best. It is thus desirable to have a data driven choice of the tuning parameters. We now introduce a practical method for selecting the tuning parameters using K -fold cross-validation.

Let $(\mathcal{X}, y) = \{(X_i, y_i), i = 1, \dots, n\}$ be the observed sample and let T be a grid of positive real values. For each $t \in T$, set

$$(\lambda, \eta) = (\lambda(t), \eta(t)) = \begin{cases} \left(t, t \left(\sqrt{\log n} (p_1 + p_2) + \sqrt{n(p_1 + p_2)} \right) \right) & \text{for ROP;} \\ \left(t, t \left(\sqrt{\log np} + \sqrt{np} \right) \right) & \text{for SROP.} \end{cases} \quad (2.41)$$

Randomly split the n samples $(X_i, y_i), i = 1, \dots, n$ into two groups of sizes $n_1 \sim \frac{(K-1)n}{K}$ and $n_2 \sim \frac{n}{K}$ for I times. Denote by $J_1^i, J_2^i \subseteq \{1, \dots, n\}$ the index sets for Group 1 and Group 2 respectively for the i -th split. Apply our procedure ((2.39) for ROP and (2.40) for SROP, respectively) to the sub-samples in Group 1 with the tuning parameters $(\lambda(t), \eta(t))$ and denote the estimators by $\hat{A}^i(t)$, $i = 1, \dots, I$. Evaluate the prediction error of $\hat{A}^i(t)$ over the sub-sample in Group 2 and set

$$\hat{R}(t) = \sum_{i=1}^I \sum_{j \in J_2^i} |y_j - \langle \hat{A}^i(t), X_j \rangle|^2, \quad t \in T.$$

We select

$$t_* = \arg \min_T \hat{R}(t)$$

and choose the tuning parameters $(\lambda(t_*), \eta(t_*))$ as in (2.41) with $t = t_*$ and the final estimator \hat{A} based on (2.39) or (2.40) with the chosen tuning parameters.

We compare the numerical result by 5-fold cross-validation with the result based on the known σ by simulation in Figure 2.5. Both the ROP and SROP are considered. It can be seen that the estimator with the tuning parameters chosen through 5-fold cross-validation has the same performance as or outperforms the one with the theoretical choice of the tuning parameters.

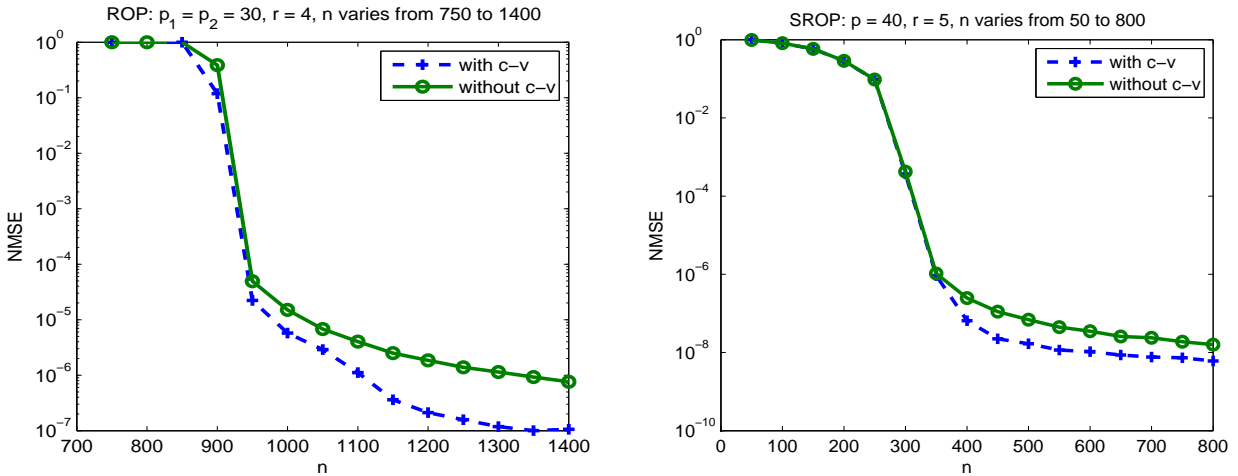


Figure 2.5: Comparison of the performance with cross validation and without cross validation in both ROP and SROP. Left panel: ROP, $p_1 = p_2 = 30, r = 4, n$ varies from 750 to 1400. Right panel: SROP, $p = 40, r = 5, n$ varies from 50 to 800.

2.5.2 Image Compression

Since a two-dimensional image can be considered as a matrix, one approach to image compression is by using low-rank matrix approximation via the singular value

decomposition. See, for example, Andrews and Patterson (1976), Recht et al. (2010), Wakin et al. (2006). Here we use an image recovery example to further illustrate the nuclear norm minimization method under the ROP model.

For a grayscale image, let $A = (a_{i,j}) \in \mathbb{R}^{m \times n}$ be the intensity matrix associated with the image, where $a_{i,j}$ is the grayscale intensity of the (i, j) pixel. When the matrix A is approximately low-rank, the ROP model and nuclear norm minimization method can be used for image compression and recovery. To illustrate this point, let us consider the following grayscale MIT Logo image (Figure 2.6).



Figure 2.6: Original grayscale MIT logo

The matrix associated with MIT logo is of the size 50×80 and of rank 6. We take rank-one random projections $\mathcal{X}(A)$ as the observed sample, with various sample sizes. Then the constrained nuclear norm minimization method is applied to reconstruct the original low-rank matrix. The recovery results are shown in Figure 2.7. The results show that the original image can be compressed and recovered well via the ROP model and the nuclear norm minimization.

2.6 Discussions

This chapter introduces the ROP model for the recovery of general low-rank matrices. A constrained nuclear norm minimization method is proposed and its theoretical and numerical properties are studied. The proposed estimator is shown to be rate-



Figure 2.7: recovery of MIT logo based on different number of measurements. Left: 900; Middle: 1000; Right: 1080.

optimal when the number of rank-one projections $n \gtrsim \log n(p_1 + p_2)$ or $n \sim r(p_1 + p_2)$. It is also shown that the procedure is adaptive to the rank and robust against small perturbations. The method and results are applied to estimation of a spiked covariance matrix. It is somewhat unexpected that it is possible to accurately recover a spiked covariance matrix from only one-dimensional projections. An interesting open problem is to estimate the principal components/subspace based on the one-dimensional random projections. We leave this as future work.

In a recent paper, Chen et al. (2013) considered quadratic measurements for the recovery of symmetric positive definite matrices, which is similar to the special case of SROP that we studied here. The paper was posted on arXiv as we finish writing the present chapter. They considered the noiseless and ℓ_1 bounded noise cases and introduced the so-called “RIP- ℓ_2/ℓ_1 ” condition. The “RIP- ℓ_2/ℓ_1 ” condition is similar to RUB in our work. But these two conditions are not identical as the RIP- ℓ_2/ℓ_1 condition can only be applied to symmetric low-rank matrices as only symmetric operators are considered in the paper. In contrast, RUB applies to all low-rank matrices.

Chen et al. (2013) (version 4) considered ℓ_1 -bounded noise case under the SROP model and gave an upper bound in their Theorem 3 (after a slight change of notation)

$$\|\hat{\Sigma} - \Sigma\|_F \leq C_1 \frac{\|\Sigma - \Sigma_\Omega\|_*}{\sqrt{r}} + C_2 \frac{\varepsilon}{n}. \quad (2.42)$$

This result for ℓ_1 bounded noise case is not applicable to the i.i.d. random noise setting. When the entries of the noise term $\eta \in \mathbb{R}^n$ are of constant order, which is the typical case for i.i.d. noise with constant variance, one has $\|\eta\|_1 \sim Cn$ with high probability. In such a case, the term $C_2 \frac{\varepsilon_1}{n}$ on the right hand side of (2.42) does not even converge to 0 as the sample size $n \rightarrow \infty$.

In comparison, the bound (2.30) in Proposition 2.3.1 can be equivalently rewritten as

$$\|\hat{A} - A\|_F \leq W_2 \frac{\|A_{-\max(r)}\|_*}{\sqrt{r}} + W_1 \tau \min \left(\frac{\sqrt{r \log n} p}{n} + \sqrt{\frac{rp}{n}}, 1 \right) \quad (2.43)$$

where the first term $W_2 \frac{\|A_{-\max(r)}\|_*}{\sqrt{r}}$ is of the same order as $C_1 \frac{\|\Sigma - \Sigma_\Omega\|_*}{\sqrt{r}}$ in (2.42) while the second term decays to 0 as $n \rightarrow \infty$. Hence, for the recovery of rank- r matrices, as the sample size n increases our bound decays to 0 but the bound (2.42) given in Chen et al. (2013) does not. The main reason of this phenomenon lies in the difference in the two methods: we use nuclear norm minimization under two convex constraints (See Remark 2.3), but Chen et al. (2013) used only the ℓ_1 constraint. Both theoretical results (see Remark 2.2.3) and numerical results (Figure 2.3 in Section 2.5) show that the additional constraint \mathcal{Z}_2 improves the performance of the estimator.

Moreover, the results and techniques in Chen et al. (2013) for symmetric positive definite matrices are not applicable to the recovery of general non-symmetric matrices. This is due to the fact that for a non-symmetric square matrix $A = (a_{ij})$, the quadratic measurements $(\beta^{(i)})^\top A \beta^{(i)}$ satisfy

$$(\beta^{(i)})^\top A \beta^{(i)} = (\beta^{(i)})^\top A^s \beta^{(i)},$$

where $A^s = \frac{1}{2}(A + A^\top)$. Hence, for a non-symmetric matrix A , only its symmetrized version A^s can be possibly identified and estimated based on the quadratic measurements, the matrix A itself is neither identifiable nor estimable.

Structured Matrix Completion

3.1 Introduction

Motivated by an array of applications, matrix completion has attracted significant recent attention in different fields including statistics, applied mathematics and electrical engineering. The central goal of matrix completion is to recover a high-dimensional low-rank matrix based on a subset of its entries. Applications include recommender systems (Koren et al., 2009), genomics (Chi et al., 2013), multi-task learning (Argyriou et al., 2008), sensor localization (Biswas et al., 2006; Singer and Cucuringu, 2010), and computer vision (Chen and Suter, 2004; Tomasi and Kanade, 1992), among many others.

Matrix completion has been well studied under the uniform sampling model, where observed entries are assumed to be sampled uniformly at random. The best known approach is perhaps the constrained nuclear norm minimization (NNM), which has been shown to yield near-optimal results when the sampling distribution of the observed entries is uniform (Candès and Recht, 2009; Candès and Tao, 2010; Gross, 2011; Recht, 2011; Candès and Plan, 2011). For estimating approximately low-rank matrices from uniformly sampled noisy observations, several penalized or constrained NNM estimators, which are based on the same principle as the well-known Lasso and

Dantzig selector for sparse signal recovery, were proposed and analyzed (Keshavan et al., 2010; Mazumder et al., 2010; Koltchinskii, 2011; Koltchinskii et al., 2011; Rohde and Tsybakov, 2011). In many applications, the entries are sampled independently but not uniformly. In such a setting, Salakhutdinov and Srebro (2010) showed that the standard NNM methods do not perform well, and proposed a weighted NNM method, which depends on the true sampling distribution. In the case of unknown sampling distribution, Foygel et al. (2011) introduced an empirically-weighted NNM method. Cai and Zhou (2013) studied a max-norm constrained minimization method for the recovery of a low-rank matrix based on the noisy observations under the non-uniform sampling model. It was shown that the max-norm constrained least squares estimator is rate-optimal under the Frobenius norm loss and yields a more stable approximate recovery guarantee with respect to the sampling distributions.

The focus of matrix completion has so far been on the recovery of a low-rank matrix based on independently sampled entries. Motivated by applications in genomic data integration, we introduce in this chapter a new framework of matrix completion called *structured matrix completion* (SMC), where a subset of the rows and a subset of the columns of an approximately low-rank matrix are observed and the goal is to reconstruct the whole matrix based on the observed rows and columns. We first discuss the genomic data integration problem before introducing the SMC model.

3.1.1 Genomic Data Integration

When analyzing genome-wide studies (GWS) of association, expression profiling or methylation, ensuring adequate power of the analysis is one of the most crucial goals due to the high dimensionality of the genomic markers under consideration. Because of cost constraints, GWS typically have small to moderate sample sizes and hence limited power. One approach to increase the power is to integrate information from

multiple GWS of the same phenotype. However, some practical complications may hamper the feasibility of such integrative analysis. Different GWS often involve different platforms with distinct genomic coverage. For example, whole genome next generation sequencing (NGS) studies would provide mutation information on all loci while older technologies for genome-wide association studies (GWAS) would only provide information on a small subset of loci. In some settings, certain studies may provide a wider range of genomic data than others. For example, one study may provide extensive genomic measurements including gene expression, miRNA and DNA methylation while other studies may only measure gene expression.

To perform integrative analysis of studies with different extent of genomic measurements, the naive complete observation only approach may suffer from low power. For the GWAS setting with a small fraction of loci missing, many imputation methods have been proposed in recent years to improve the power of the studies. Examples of useful methods include haplotype reconstruction, k -nearest neighbor, regression and singular value decomposition methods (Scheet and Stephens, 2006; Li and Abecasis, 2006; Browning and Browning, 2009; Troyanskaya et al., 2001; Kim et al., 2005; Wang et al., 2006). Many of the haplotype phasing methods are considered to be highly effective in recovering missing genotype information (Yu and Schaid, 2007). These methods, while useful, are often computationally intensive. In addition, when one study has a much denser coverage than the other, the fraction of missingness could be high and an exceedingly large number of observation would need to be imputed. It is unclear whether it is statistically or computationally feasible to extend these methods to such settings. Moreover, haplotype based methods cannot be extended to incorporate other types of genomic data such as gene expression and miRNA data.

When integrating multiple studies with different extent of genomic measurements, the observed data can be viewed as complete rows and columns of a large matrix

A and the missing components can be arranged as a submatrix of A . As such, the missingness in A is structured by design. In this chapter, we propose a novel SMC method for imputing the missing submatrix of A . As shown in Section 3.5, by imputing the missing miRNA measurements and constructing prediction rules based on the imputed data, it is possible to significantly improve the prediction performance.

3.1.2 Structured Matrix Completion Model

Motivated by the applications mentioned above, this chapter considers SMC where a subset of rows and columns are observed. Specifically, we observe $m_1 < p_1$ rows and $m_2 < p_2$ columns of a matrix $A \in \mathbb{R}^{p_1 \times p_2}$ and the goal is to recover the whole matrix. Since the singular values are invariant under row/column permutations, it can be assumed without loss of generality that we observe the first m_1 rows and m_2 columns of A which can be written in a block form:

$$A = \begin{matrix} & \begin{matrix} m_2 & p_2 - m_2 \end{matrix} \\ \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} & \begin{matrix} m_1 \\ p_1 - m_1 \end{matrix} \end{matrix} \quad (3.1)$$

where A_{11} , A_{12} , and A_{21} are observed and the goal is to recover the missing block A_{22} . See Figure 3.1(a) in Section 3.2 for a graphical display of the data. Clearly there is no way to recover A_{22} if A is an arbitrary matrix. However, in many applications such as genomic data integration discussed earlier, A is approximately low-rank, which makes it possible to recover A_{22} with accuracy. In this chapter, we introduce a method based on the singular value decomposition (SVD) for the recovery of A_{22} when A is approximately low-rank.

It is important to note that the observations here are much more “structured” comparing to the previous settings of matrix completion. As the observed entries are

in full rows or full columns, the existing methods based on NNM are not suitable. As mentioned earlier, constrained NNM methods have been widely used in matrix completion problems based on independently observed entries. However, for the problem considered in the present chapter, these methods do not utilize the structure of the observations and do not guarantee precise recovery even for exactly low-rank matrix A (See Remark 3.2.1 in Section 3.2). Numerical results in Section 3.4 show that NNM methods do not perform well in SMC.

In this chapter we propose a new SMC method that can be easily implemented by a fast algorithm which only involves basic matrix operations and the SVD. The main idea of our recovery procedure is based on the Schur Complement. In the ideal case when A is exactly low rank, the Schur complement of the missing block, $A_{22} - A_{21}A_{11}^\dagger A_{12}$, is zero and thus $A_{21}A_{11}^\dagger A_{12}$ can be used to recover A_{22} exactly. When A is approximately low rank, $A_{21}A_{11}^\dagger A_{12}$ cannot be used directly to estimate A_{22} . For this case, we transform the observed blocks using SVD; remove some unimportant rows and columns based on thresholding rules; and subsequently apply a similar procedure to recover A_{22} .

Both its theoretical and numerical properties are studied. It is shown that the estimator recovers low-rank matrices accurately and is robust against small perturbations. A lower bound result shows that the estimator is rate optimal for a class of approximately low-rank matrices. Although it is required for the theoretical analysis that there is a significant gap between the singular values of the true low-rank matrix and those of the perturbation, simulation results indicate that this gap is not really necessary in practice and the estimator recovers A accurately whenever the singular values of A decay sufficiently fast.

3.1.3 Organization of the Paper

The rest of the chapter is organized as follows. In Section 3.2, we introduce in detail the proposed SMC methods when A is exactly or approximately low-rank. The theoretical properties of the estimators are analyzed in Section 3.3. Both upper and lower bounds for the recovery accuracy under the Schatten- q norm loss are established. Simulation results are shown in Section 3.4 to investigate the numerical performance of the proposed methods. A real data application to genomic data integration is given in Section 3.5. Section 3.6 discusses a few practical issues related to real data applications. For reasons of space, the proofs of the main results and additional simulation results are given in the Appendix. Some key technical tools used in the proofs of the main theorems are also developed and proved in the Appendix.

3.2 Structured Matrix Completion: Methodology

In this section, we propose procedures to recover the submatrix A_{22} based on the observed blocks A_{11} , A_{12} , and A_{21} . We begin with basic notation and definitions that will be used in the rest of the chapter.

For a matrix U , we use $U_{[\Omega_1, \Omega_2]}$ to represent its sub-matrix with row indices Ω_1 and column indices Ω_2 . We also use the Matlab syntax to represent index sets. Specifically for integers $a \leq b$, “ $a : b$ ” represents $\{a, a + 1, \dots, b\}$; and “ $:$ ” alone represents the entire index set. Therefore, $U_{[:, 1:r]}$ stands for the first r columns of U while $U_{[(m_1+1):p_1, :]}$ stands for the $\{m_1 + 1, \dots, p_1\}^{th}$ rows of U . For the matrix A given in (3.1), we use the notation $A_{\bullet 1}$ and $A_{1\bullet}$ to denote $[A_{11}^\top, A_{21}^\top]^\top$ and $[A_{11}, A_{12}]$, respectively. For a matrix $B \in \mathbb{R}^{m \times n}$, let $B = U\Sigma V^\top = \sum_i \sigma_i(B) u_i v_i^\top$ be the SVD, where $\Sigma = \text{diag}\{\sigma_1(B), \sigma_2(B), \dots\}$ with $\sigma_1(B) \geq \sigma_2(B) \geq \dots \geq 0$ being the singular values of B in decreasing order. The smallest singular value $\sigma_{\min(m,n)}$, which will

be denoted by $\sigma_{\min}(B)$, plays an important role in our analysis. We also define $B_{\max(r)} = \sum_{i=1}^r \sigma_i(B) u_i v_i^\top$ and $B_{-\max(r)} = B - B_{\max(r)} = \sum_{i \geq r+1} \sigma_i(B) u_i v_i^\top$. For $1 \leq q \leq \infty$, the Schatten- q norm $\|B\|_q$ is defined to be the vector q -norm of the singular values of B , i.e. $\|B\|_q = (\sum_i \sigma_i^q(B))^{1/q}$. Three special cases are of particular interest: when $q = 1$, $\|B\|_1 = \sum_i \sigma_i(B)$ is the nuclear (or trace) norm of B and will be denoted as $\|B\|_*$; when $q = 2$, $\|B\|_2 = \sqrt{\sum_{i,j} B_{ij}^2}$ is the Frobenius norm of B and will be denoted as $\|B\|_F$; when $q = \infty$, $\|B\|_\infty = \sigma_1(B)$ is the spectral norm of B that we simply denote as $\|B\|$. For any matrix $U \in \mathbb{R}^{p \times n}$, we use $P_U \equiv U(U^\top U)^\dagger U^\top \in \mathbb{R}^{p \times p}$ to denote the projection operator onto the column space of U . Throughout, we assume that A is *approximately rank r* in that for some integer $0 < r \leq \min(m_1, m_2)$, there is a significant gap between $\sigma_r(A)$ and $\sigma_{r+1}(A)$ and the tail $\|A_{-\max(r)}\|_q = (\sum_{k \geq r+1} \sigma_k^q(A))^{1/q}$ is small. The gap assumption enables us to provide a theoretical upper bound on the accuracy of the estimator, while it is not necessary in practice (see Section 3.4 for more details).

3.2.1 Exact Low-rank Matrix Recovery

We begin with the relatively easy case where A is exactly of rank r . In this case, a simple analysis indicates that A can be perfectly recovered as shown in the following proposition.

Proposition 3.2.1. *Suppose A is of rank r , the SVD of A_{11} is $A_{11} = U\Sigma V^\top$, where $U \in \mathbb{R}^{p_1 \times r}$, $\Sigma \in \mathbb{R}^{r \times r}$, and $V \in \mathbb{R}^{p_2 \times r}$. If*

$$\text{rank}([A_{11} \ A_{12}]) = \text{rank} \left(\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \right) = \text{rank}(A) = r,$$

then $\text{rank}(A_{11}) = r$ and A_{22} is exactly given by

$$A_{22} = A_{21}(A_{11})^\dagger A_{12} = A_{21}V(\Sigma)^{-1}U^\top A_{12}. \quad (3.2)$$

Remark 3.2.1. Under the same conditions as Proposition 3.2.1, the NNM

$$\hat{A}_{22} = \arg \min_B \left\| \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & B \end{bmatrix} \right\|_* \quad (3.3)$$

fails to guarantee the exact recovery of A_{22} . Consider the case where A is a $p_1 \times p_2$ matrix with all entries being 1. Suppose we observe arbitrary m_1 rows and m_2 columns, the NNM would yield $\hat{A}_{22} \in \mathbb{R}^{(p_1-m_1) \times (p_2-m_2)}$ with all entries being $\left(1 \wedge \sqrt{\frac{m_1 m_2}{(p_1-m_1)(p_2-m_2)}}\right)$ (See Lemma A.3.4 in the Appendix). Hence when $m_1 m_2 < (p_1 - m_1)(p_2 - m_2)$, i.e., when the size of the observed blocks are much smaller than that of A , the NNM fails to recover exactly the missing block A_{22} . See also the numerical comparison in Section 3.4. The NNM (3.3) also fails to recover A_{22} with high probability in a random matrix setting where $A = B_1 B_2^\top$ with $B_1 \in \mathbb{R}^{p_1 \times r}$ and $B_2 \in \mathbb{R}^{p_2 \times r}$ being i.i.d. standard Gaussian matrices. See Lemma A.3.3 in the Appendix for further details. In addition to (3.3), other variations of NNM have been proposed in the literature, including *penalized* NNM (Toh and Yun, 2010; Mazumder et al., 2010),

$$\hat{A}^{PN} = \arg \min_Z \left\{ \frac{1}{2} \sum_{(i_k, j_k) \in \Omega} (Z_{i_k, j_k} - A_{i_k, j_k})^2 + t \|Z\|_* \right\}; \quad (3.4)$$

and *constrained* NNM with relaxation (Cai et al., 2010a),

$$\hat{A}^{CN} = \arg \min_Z \{ \|Z\|_* : |Z_{i_k, j_k} - A_{i_k, j_k}| \leq t \text{ for } (i_k, j_k) \in \Omega \}, \quad (3.5)$$

where $\Omega = \{(i_k, j_k) : A_{i_k, j_k} \text{ observed}, 1 \leq i_k \leq p_1, 1 \leq j_k \leq p_2\}$ and t is the tuning parameter. However, these NNM methods may not be suitable for SMC especially when only a small number of rows and columns are observed. In particular, when $m_1 \ll p_1, m_2 \ll p_2$, A is well spread in each block $A_{11}, A_{12}, A_{21}, A_{22}$, we have $\| [A_{11} \ A_{12}] \|_* \ll \|A\|_*, [A_{12}]_* \ll \|A\|_*$. Thus,

$$\left\| \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & 0 \end{bmatrix} \right\|_* \leq \left\| \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \right\|_* + \left\| [A_{12}] \right\|_* \ll \left\| \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \right\|_*.$$

In the other words, imputing A_{22} with all zero yields a much smaller nuclear norm than imputing with the true A_{22} and hence NNM methods would generally fail to recover A_{22} under such settings.

Proposition 3.2.1 shows that, when A is exactly low-rank, A_{22} can be recovered precisely by $A_{21}(A_{11})^\dagger A_{12}$. Unfortunately, this result heavily relies on the exactly low-rank assumption that cannot be directly used for approximately low-rank matrices. In fact, even with a small perturbation to A , the inverse of A_{11} makes the formula $A_{21}(A_{11})^\dagger A_{12}$ unstable, which may lead to the failure of recovery. In practice, A is often not exactly low rank but approximately low rank. Thus for the rest of the chapter, we focus on the latter setting.

3.2.2 Approximate Low-rank Matrix Recovery

Let $A = U\Sigma V^\top$ be the SVD of an approximately low rank matrix A and partition $U \in \mathbb{R}^{p_1 \times p_1}$, $V \in \mathbb{R}^{p_2 \times p_2}$ and $\Sigma \in \mathbb{R}^{p_1 \times p_2}$ into blocks as

$$U = \begin{bmatrix} r & p_1 - r \\ U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \begin{matrix} m_1 \\ p_1 - m_1 \end{matrix}, \quad V = \begin{bmatrix} r & p_2 - r \\ V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \begin{matrix} m_2 \\ p_2 - m_2 \end{matrix}, \quad \Sigma = \begin{bmatrix} r & p_2 - r \\ \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{matrix} r \\ p_1 - r \end{matrix} \quad (3.6)$$

Then A can be decomposed as $A = A_{\max(r)} + A_{-\max(r)}$ where $A_{\max(r)}$ is of rank r with the largest r singular values of A and $A_{-\max(r)}$ is general but with small singular values. Then

$$A_{\max(r)} = U_{\bullet 1} \Sigma_1 V_{\bullet 1}^\top = \begin{bmatrix} m_2 & p_2 - m_2 \\ U_{11} \Sigma_1 V_{11}^\top & U_{11} \Sigma_1 V_{21}^\top \\ U_{21} \Sigma_1 V_{11}^\top & U_{21} \Sigma_1 V_{21}^\top \end{bmatrix} \begin{matrix} m_1 \\ p_1 - m_1 \end{matrix}, \quad \text{and} \quad A_{-\max(r)} = U_{\bullet 2} \Sigma_2 V_{\bullet 2}^\top. \quad (3.7)$$

Here and in the sequel, we use the notation $U_{\bullet k}$ and $U_{k\bullet}$ to denote $[U_{1k}^\top, U_{2k}^\top]^\top$ and $[U_{k1}, U_{k2}]$, respectively. Thus, $A_{\max(r)}$ can be viewed as a rank- r approximation to A and obviously

$$U_{21} \Sigma_1 V_{21}^\top = \{U_{21} \Sigma_1 V_{11}^\top\} \{U_{11} \Sigma_1 V_{11}^\top\}^{-1} \{U_{11} \Sigma_1 V_{21}^\top\}.$$

We will use the observed A_{11} , A_{12} and A_{21} to obtain estimates of $U_{\bullet 1}$, $V_{\bullet 1}$ and Σ_1 and subsequently recover A_{22} using an estimated $U_{21} \Sigma_1 V_{21}^\top$.

When r is known, i.e., we know where the gap is located in the singular values of A , a simple procedure can be implemented to estimate A_{22} as described in Algorithm

1 below by estimating $U_{\bullet 1}$ and $V_{\bullet 1}$ using the principal components of $A_{\bullet 1}$ and $A_{1\bullet}$.

Algorithm 1 Algorithm for Structured Matrix Completion with a given r

- 1: Input: $A_{11} \in \mathbb{R}^{m_1 \times m_2}$, $A_{12} \in \mathbb{R}^{(p_1 - m_1) \times m_2}$, $A_{21} \in \mathbb{R}^{m_1 \times (p_2 - m_2)}$.
- 2: Calculate the SVD of $A_{\bullet 1}$ and $A_{1\bullet}$ to obtain $A_{\bullet 1} = U^{(1)}\Sigma^{(1)}V^{(1)\top}$, $A_{1\bullet} = U^{(2)}\Sigma^{(2)}V^{(2)\top}$.
- 3: Suppose M, N are orthonormal basis of U_{11}, V_{11} . We estimate the column space of U_{11} and V_{11} by $\hat{M} = U_{[:,1:r]}^{(2)}$, $\hat{N} = V_{[:,1:r]}^{(1)}$.
- 4: Finally we estimate A_{22} as

$$\hat{A}_{22} = A_{21}\hat{N}(\hat{M}^\top A_{11}\hat{N})^{-1}\hat{M}^\top A_{12}. \quad (3.8)$$

However, Algorithm 1 has several major limitations. First, it relies on a given r which is typically unknown in practice. Second, the algorithm need to calculate the matrix divisions, which may cause serious precision issues when the matrix is near-singular or the rank r is mis-specified. To overcome these difficulties, we propose another Algorithm which essentially first estimates r with \hat{r} and then apply Algorithm 1 to recover A_{22} . Before introducing the algorithm of recovery without knowing r , it is helpful to illustrate the idea with heat maps in Figures 3.1 and 3.2.

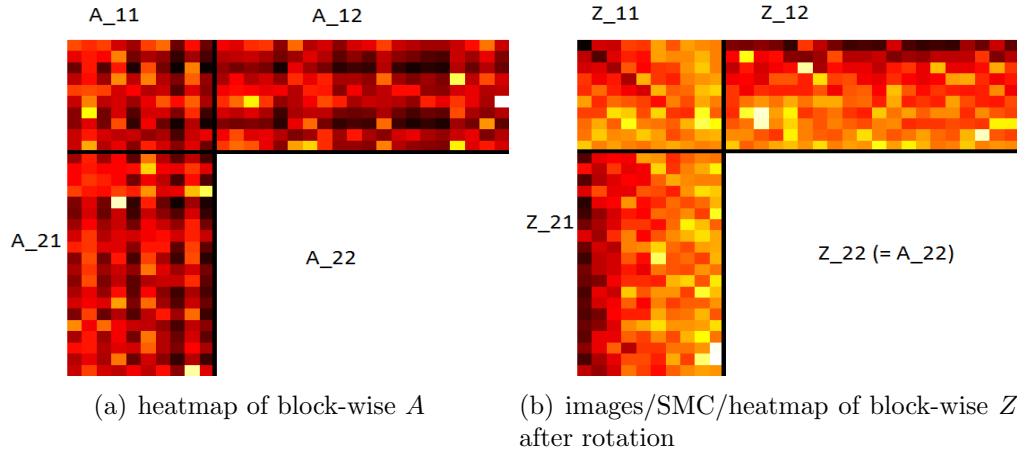


Figure 3.1: Illustrative example with $A \in \mathbb{R}^{30 \times 30}$, $m_1 = m_2 = 10$. (A darker block corresponds to larger magnitude.)

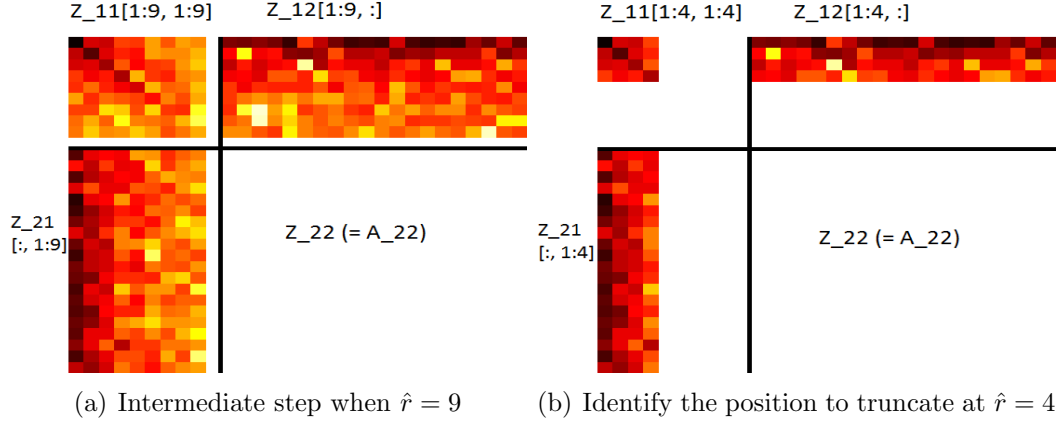


Figure 3.2: Searching for the appropriate position to truncate from $\hat{r} = 10$ to 1.

Our procedure has three steps.

1. First, we move the significant factors of $A_{\bullet 1}$ and $A_{1\bullet}$ to the front by rotating the columns of $A_{\bullet 1}$ and the rows of $A_{1\bullet}$ based on the SVD,

$$A_{\bullet 1} = U^{(1)}\Sigma^{(1)}V^{(1)\top}, \quad A_{1\bullet} = U^{(2)}\Sigma^{(2)}V^{(2)\top}.$$

After the transformation, we have Z_{11}, Z_{12}, Z_{21} ,

$$Z_{11} = U^{(2)\top}A_{11}V^{(1)}, \quad Z_{12} = U^{(2)\top}A_{12}, \quad Z_{21} = A_{21}V^{(1)}, \quad Z_{22} = A_{22}.$$

Clearly A and Z have the same singular values since the transformation is orthogonal. As shown in Figure 3.1(b), the amplitudes of the columns of $Z_{\bullet 1} = [Z_{11}^\top, Z_{21}^\top]^\top$ and the rows of $Z_{1\bullet} = [Z_{11}, Z_{12}]$ are decaying.

2. When A is exactly of rank r , the $\{r + 1, \dots, m_1\}^{th}$ rows and $\{r + 1, \dots, m_2\}^{th}$ columns of Z are zero. Due to the small perturbation term $A_{-\max(r)}$, the back columns of $Z_{\bullet 1}$ and rows of $Z_{1\bullet}$ are small but non-zero. In order to recover $A_{\max(r)}$, the best rank r approximation to A , a natural idea is to first delete

these back rows of $Z_{1\bullet}$ and columns of $Z_{\bullet 1}$, i.e. the $\{r + 1, \dots, m_1\}^{th}$ rows and $\{r + 1, \dots, m_2\}^{th}$ columns of Z .

However, since r is unknown, it is unclear how many back rows and columns should be removed. It will be helpful to have an estimate for r , \hat{r} , and then use $Z_{21,[:,1:\hat{r}]}$, $Z_{11,[1:\hat{r},1:\hat{r}]}$ and $Z_{12[1:\hat{r},:]}$ to recover A_{22} . It will be shown that a good choice of \hat{r} would satisfy that $Z_{11,[1:\hat{r},1:\hat{r}]}$ is non-singular and $\|Z_{21,[1:\hat{r},1:\hat{r}]}Z_{11,[1:\hat{r},1:\hat{r}]}^{-1}\| \leq T_R$, where T_R is some constant to be specified later. Our final estimator for r would be the largest \hat{r} that satisfies this condition, which can be identified recursively from $\min(m_1, m_2)$ to 1 (See Figure 3.2).

3. Finally, similar to (3.2), A_{22} can be estimated by

$$\hat{A}_{22} = Z_{21,[:,1:\hat{r}]}Z_{11,[1:\hat{r},1:\hat{r}]}^{-1}Z_{12,[1:\hat{r},:]}, \quad (3.9)$$

The method we propose can be summarized as the following algorithm.

Algorithm 2 Algorithm of Structured Matrix Completion with unknown r

- 1: Input: $A_{11} \in \mathbb{R}^{m_1 \times m_2}$, $A_{12}^{m_1 \times (p_2 - m_2)}$, $A_{21}^{(p_1 - m_1) \times m_2}$. Thresholding level: T_R , (or T_C).
- 2: Calculate the SVD $A_{\bullet 1} = U^{(1)}\Sigma^{(1)}V^{(1)\top}$, $A_{1\bullet} = U^{(2)}\Sigma^{(2)}V^{(2)\top}$.
- 3: Calculate $Z_{11} \in \mathbb{R}^{m_1 \times m_2}$, $Z_{12} \in \mathbb{R}^{m_1 \times (p_2 - m_2)}$, $Z_{21} \in \mathbb{R}^{(p_1 - m_1) \times m_2}$

$$Z_{11} = U^{(2)\top} A_{11} V^{(1)}, \quad Z_{12} = U^{(2)\top} A_{12}, \quad Z_{21} = A_{21} V^{(1)}.$$

- 4: **for** $s = \min(m_1, m_2) : -1 : 1$ **do** (Use iteration to find \hat{r})
- 5: Calculate $D_{R,s} \in \mathbb{R}^{(p_1 - m_1) \times s}$ (or $D_{C,s} \in \mathbb{R}^{s \times (p_2 - m_2)}$) by solving linear equation system,

$$D_{R,s} = Z_{21,[:,1:s]} Z_{11,[:,1:s]}^{-1} \quad (\text{or} \quad D_{C,s} = Z_{11,[:,1:s]}^{-1} Z_{12,[:,1:s]})$$

- 6: **if** $Z_{11,[:,1:s]}$ is not singular and $\|D_{R,s}\| \leq T_R$ (or $\|D_{C,s}\| \leq T_C$) **then**
- 7: $\hat{r} = s$; **break** from the loop;
- 8: **end if**
- 9: **end for**
- 10: **if** (\hat{r} is not valued) **then** $\hat{r} = 0$.
- 11: **end if**
- 12: Finally we calculate the estimate as

$$\hat{A}_{22} = Z_{21,[:,1:\hat{r}]} Z_{11,[:,1:\hat{r}]}^{-1} Z_{12,[:,1:\hat{r}]}$$

It can also be seen from Algorithm 2 that the estimator \hat{r} is constructed based on either the row thresholding rule $\|D_{R,s}\| \leq T_R$ or the column thresholding rule $\|D_{C,s}\| \leq T_C$. Discussions on the choice between $D_{R,s}$ and $D_{C,s}$ are given in the next section. Let us focus for now on the row thresholding based on $D_{R,s} = Z_{21,[:,1:s]} Z_{11,[:,1:s]}^{-1}$. It is important to note that $Z_{21[:,1:r]}$ and $Z_{11,[:,1:r]}$ approximate $U_{21}\Sigma_1$ and Σ_1 , respectively. The idea behind the proposed \hat{r} is that when $s > r$, $Z_{21[:,1:s]}$ and $Z_{11,[:,1:s]}$ are nearly singular and hence $D_{R,s}$ may either be deemed singular or with unbounded norm. When $s = r$, $Z_{11,[:,1:s]}$ is non-singular with $\|D_{R,s}\|$ bounded by some constant, as we show in Theorem 3.3.2. Thus, we estimate \hat{r} as the largest r such that $Z_{11,[:,1:s]}$ is non-singular with $\|D_{R,s}\| < T_R$.

3.3 Theoretical Analysis

In this section, we investigate the theoretical properties of the algorithms introduced in Section 3.2. Upper bounds for the estimation errors of Algorithms 1 and 2 are presented in Theorems 3.3.1 and 3.3.2, respectively, and the lower-bound results are given in Theorem 3.3.3. These bounds together establish the optimal rate of recovery over certain classes of approximately low-rank matrices. The choices of tuning parameters T_R and T_C are discussed in Corollaries 3.3.1 and 3.3.2.

Theorem 3.3.1. *Suppose \hat{A} is given by the procedure of Algorithm 1. Assume*

$$\sigma_{r+1}(A) \leq \frac{1}{2} \sigma_r(A) \cdot \sigma_{\min}(U_{11}) \cdot \sigma_{\min}(V_{11}), \quad (3.10)$$

Then for any $1 \leq q \leq \infty$,

$$\left\| \hat{A}_{22} - A_{22} \right\|_q \leq 3 \|A_{-\max(r)}\|_q \left(1 + \frac{1}{\sigma_{\min}(U_{11})} \right) \left(1 + \frac{1}{\sigma_{\min}(V_{11})} \right) \quad (3.11)$$

Remark 3.3.1. It is helpful to explain intuitively why Condition (3.10) is needed. When A is approximately low-rank, the dominant low-rank component of A , $A_{\max(r)}$, serves as a good approximation to A , while the residual $A_{-\max(r)}$ is “small”. The goal is to recover $A_{\max(r)}$ well. Among the three observed blocks, A_{11} is the most important and it is necessary to have $A_{\max(r)}$ dominating $A_{-\max(r)}$ in A_{11} . Note that $A_{11} = A_{\max(r), [1:m_1, 1:m_2]} + A_{-\max(r), [1:m_1, 1:m_2]}$,

$$\sigma_r(A_{\max(r), [1:m_1, 1:m_2]}) = \sigma_r(U_{11} \Sigma_1 V_{11}^T) \geq \sigma_{\min}(U_{11}) \sigma_r(A) \sigma_{\min}(V_{11}),$$

$$\|A_{-\max(r), [1:m_1, 1:m_2]}\| = \|U_{12} \Sigma_2 V_{12}^T\| \leq \sigma_{r+1}(A).$$

We thus require Condition (3.10) in Theorem 3.3.1 for the theoretical analysis.

Theorem 3.3.1 gives an upper bound for the estimation accuracy of Algorithm 1 under the assumption that there is a significant gap between $\sigma_r(A)$ and $\sigma_{r+1}(A)$ for some known r . It is noteworthy that there are possibly multiple values of r that satisfy Condition (3.10). In such a case, the bound (3.11) applies to all such r and the largest r yields the strongest result.

We now turn to Algorithm 2, where the knowledge of r is not assumed. Theorem 3.3.2 below shows that for properly chosen T_R or T_C , Algorithm 2 can lead to accurate recovery of A_{22} .

Theorem 3.3.2. *Assume that there exists $r \in [1, \min(m_1, m_2)]$ such that*

$$\sigma_{r+1}(A) \leq \frac{1}{4} \sigma_r(A) \cdot \sigma_{\min}(U_{11}) \sigma_{\min}(V_{11}). \quad (3.12)$$

Let T_R and T_C be two constants satisfying

$$T_R \geq \frac{1.36}{\sigma_{\min}(U_{11})} + 0.35 \quad \text{and} \quad T_C \geq \frac{1.36}{\sigma_{\min}(V_{11})} + 0.35.$$

Then for $1 \leq q \leq \infty$, \hat{A}_{22} given by Algorithm 2 satisfies

$$\begin{aligned} \left\| \hat{A}_{22} - A_{22} \right\|_q &\leq 6.5 T_R \left(\frac{1}{\sigma_{\min}(V_{11})} + 1 \right) \|A_{-\max(r)}\|_q \\ \text{or} \quad \left\| \hat{A}_{22} - A_{22} \right\|_q &\leq 6.5 T_C \left(\frac{1}{\sigma_{\min}(U_{11})} + 1 \right) \|A_{-\max(r)}\|_q \end{aligned} \quad (3.13)$$

when \hat{r} is estimated based on the thresholding rule $\|D_{R,s}\| \leq T_R$ or $\|D_{C,s}\| \leq T_C$, respectively.

Besides $\sigma_r(A)$ and $\sigma_{r+1}(A)$, Theorems 3.3.1 and 3.3.2 involve $\sigma_{\min}(U_{11})$ and $\sigma_{\min}(V_{11})$, two important quantities that reflect how much the low-rank matrix $A_{\max(r)} = U_{\bullet 1} \Sigma_1 V_{\bullet 1}^\top$ is concentrated on the first m_1 rows and m_2 columns. We should note that $\sigma_{\min}(U_{11})$ and $\sigma_{\min}(V_{11})$ depend on the singular vectors of A and $\sigma_r(A)$ and $\sigma_{r+1}(A)$ are the sin-

gular values of A . The lower bound in Theorem 3.3.3 below indicates that $\sigma_{\min}(U_{11})$, $\sigma_{\min}(V_{11})$, and the singular values of A together quantify the difficulty of the problem: recovery of A_{22} gets harder as $\sigma_{\min}(U_{11})$ and $\sigma_{\min}(V_{11})$ become smaller or the $\{r + 1, \dots, \min(p_1, p_2)\}^{th}$ singular values become larger. Define the class of approximately rank- r matrices $\mathcal{F}_r(M_1, M_2)$ by

$$\mathcal{F}_r(M_1, M_2) = \left\{ A \in \mathbb{R}^{p_1 \times p_2} : \begin{array}{l} \sigma_{\min}(U_{11}) \geq M_1, \sigma_{\min}(V_{11}) \geq M_2, \\ \sigma_{r+1}(A) \leq \frac{1}{2}\sigma_r(A)\sigma_{\min}(U_{11})\sigma_{\min}(V_{11}) \end{array} \right\}. \quad (3.14)$$

Theorem 3.3.3 (Lower Bound). *Suppose $r \leq \min(m_1, m_2, p_1 - m_1, p_2 - m_2)$ and $0 < M_1, M_2 < 1$, then for all $1 \leq q \leq \infty$,*

$$\inf_{\hat{A}_{22}} \sup_{A \in \mathcal{F}_r(M_1, M_2)} \frac{\|\hat{A}_{22} - A_{22}\|_q}{\|A_{-\max(r)}\|_q} \geq \frac{1}{4} \left(\frac{1}{M_1} + 1 \right) \left(\frac{1}{M_2} + 1 \right). \quad (3.15)$$

Remark 3.3.2. Theorems 3.3.1, 3.3.2 and 3.3.3 together immediately yield the optimal rate of recovery over the class $\mathcal{F}_r(M_1, M_2)$,

$$\begin{aligned} \inf_{\hat{A}_{22}} \sup_{A \in \mathcal{F}_r(M_1, M_2)} \frac{\|\hat{A}_{22} - A_{22}\|_q}{\|A_{-\max(r)}\|_q} \\ \asymp \left(\frac{1}{M_1} + 1 \right) \left(\frac{1}{M_2} + 1 \right) \text{ for } 0 \leq M_1, M_2 < 1, 1 \leq q \leq \infty. \end{aligned} \quad (3.16)$$

Since U_{11} and V_{11} are determined by the SVD of A and $\sigma_{\min}(U_{11})$ and $\sigma_{\min}(V_{11})$ are unknown based only on A_{11} , A_{12} , and A_{21} , it is thus not straightforward to choose the tuning parameters T_R and T_C in a principled way. Theorem 3.3.2 also does not provide information on the choice between row and column thresholding. Such a choice generally depends on the problem setting. We consider below two settings where either the row/columns of A are randomly sampled or A is itself a random low-rank matrix. In such settings, when A is approximately rank r and at least $O(r \log r)$

number of rows and columns are observed, Algorithm 2 gives accurate recovery of A with fully specified tuning parameter. We first consider in Corollary 3.3.1 a fixed matrix A with the observed m_1 rows and m_2 columns selected uniformly randomly.

Corollary 3.3.1 (Random Rows/Columns). *Let $A = U\Sigma V^\top$ be the SVD of $A \in \mathbb{R}^{p_1 \times p_2}$. Set*

$$W_r^{(1)} = \frac{p_1}{r} \max_{1 \leq i \leq p_1} \sum_{j=1}^r U_{ij}^2 \quad \text{and} \quad W_r^{(2)} = \frac{p_2}{r} \max_{1 \leq i \leq p_2} \sum_{j=1}^r V_{ij}^2. \quad (3.17)$$

Let $\Omega_1 \subset \{1, \dots, p_1\}$ and $\Omega_2 \subset \{1, \dots, p_2\}$ be respectively the index set of the observed m_1 rows and m_2 columns. Then A can be decomposed as

$$A_{11} = A_{[\Omega_1, \Omega_2]}, \quad A_{21} = A_{[\Omega_1^c, \Omega_2]}, \quad A_{12} = A_{[\Omega_1, \Omega_2^c]}, \quad A_{22} = A_{[\Omega_1^c, \Omega_2^c]}. \quad (3.18)$$

1. Let Ω_1 and Ω_2 be independently and uniformly selected from $\{1, \dots, p_1\}$ and $\{1, \dots, p_2\}$ with or without replacement, respectively. Suppose there exists $r \leq \min(m_1, m_2)$ such that

$$\sigma_{r+1}(A) \leq \frac{1}{6} \sigma_r(A) \sqrt{\frac{m_1 m_2}{p_1 p_2}}.$$

and the number of rows and number of columns we observed satisfy

$$m_1 \geq 12.5rW_r^{(1)}(\log(r) + c), \quad m_2 \geq 12.5rW_r^{(2)}(\log(r) + c),$$

for some constant $c > 1$. Algorithm 2 with either column thresholding with the break condition $\|D_{R,s}\| \leq T_R$ where $T_R = 2\sqrt{\frac{p_1}{m_1}}$ or row thresholding with the

break condition $\|D_{C,s}\| \leq T_C$ where $T_C = 2\sqrt{\frac{p_2}{m_2}}$ satisfies, for all $1 \leq q \leq \infty$,

$$\|\hat{A}_{22} - A_{22}\|_q \leq 29\|A_{-\max(r)}\|_q \sqrt{\frac{p_1 p_2}{m_1 m_2}} \quad \text{with probability } \geq 1 - 4\exp(-c).$$

2. If Ω_1 is uniformly randomly selected from $\{1, \dots, p_1\}$ with or without replacement (Ω_2 is not necessarily random), and there exists $r \leq m_2$ such that

$$\sigma_{r+1}(A) \leq \frac{1}{5}\sigma_r(A)\sigma_{\min}(V_{11})\sqrt{\frac{m_1}{p_1}}$$

and the number of observed rows satisfies

$$m_1 \geq 12.5rW_r^{(1)}(\log(r) + c) \quad \text{for some constant } c > 1, \quad (3.19)$$

then Algorithm 2 with the break condition $\|D_{R,s}\| \leq T_R$ where $T_R \geq 2\sqrt{\frac{p_1}{m_1}}$ satisfies, for all $1 \leq q \leq \infty$,

$$\|\hat{A}_{22} - A_{22}\|_q \leq 6.5\|A_{-\max(r)}\|_q T_R \left(\frac{1}{\sigma_{\min}(V_{11})} + 1 \right)$$

with probability at least $1 - 2\exp(-c)$.

3. Similarly, if Ω_2 is uniformly randomly selected from $\{1, \dots, p_2\}$ with or without replacement (Ω_1 is not necessarily random) and there exists $r \leq m_2$ such that

$$\sigma_{r+1}(A) \leq \frac{1}{5}\sigma_r(A)\sigma_{\min}(U_{11})\sqrt{\frac{m_2}{p_2}},$$

and the number of observed columns satisfies

$$m_2 \geq 12.5rW_r^{(2)}(\log(r) + c) \quad \text{for some constant } c > 1, \quad (3.20)$$

then Algorithm 2 with the break condition $\|D_{C,s}\| \leq T_C$ where $T_C \geq 2\sqrt{\frac{p_2}{m_2}}$ satisfies, for all $1 \leq q \leq \infty$,

$$\left\| \hat{A}_{22} - A_{22} \right\|_q \leq 6.5 \|A_{-\max(r)}\|_q T_C \left(\frac{1}{\sigma_{\min}(U_{11})} + 1 \right)$$

with probability at least $1 - 2\exp(-c)$.

Remark 3.3.3. The quantities $W_r^{(1)}$ and $W_r^{(2)}$ in Corollary 3.3.1 measure the variation of amplitude of each row or each column of $A_{\max(r)}$. When $W_r^{(1)}$ and $W_r^{(2)}$ become larger, a small number of rows and columns in $A_{\max(r)}$ would have larger amplitude than others, while these rows and columns would be missed with large probability in the sampling of Ω , which means the problem would become harder. Hence, more observations for the matrix with larger $W_r^{(1)}$ and $W_r^{(2)}$ are needed as shown in (3.19).

We now consider the case where A is a random matrix.

Corollary 3.3.2 (Random Matrix). *Suppose $A \in \mathbb{R}^{p_1 \times p_2}$ is a random matrix generated by $A = U\Sigma V^\top$, where the singular values Σ and singular space V are fixed, and U has orthonormal columns that are randomly sampled based on the Haar measure. Suppose we observe the first m_1 rows and first m_2 columns of A . Assume there exists $r < \frac{1}{2} \min(m_1, m_2)$ such that*

$$\sigma_{r+1}(A) \leq \frac{1}{5} \sigma_r(A) \sigma_{\min}(V_{11}) \sqrt{\frac{m_1}{p_1}}.$$

Then there exist uniform constants $c, \delta > 0$ such that if $m_1 \geq cr$, \hat{A}_{22} is given by Algorithm 2 with the break condition $\|D_{R,s}\| \leq T_R$, where $T_R \geq 2\sqrt{\frac{p_1}{m_1}}$, we have for all $1 \leq q \leq \infty$,

$$\left\| \hat{A}_{22} - A_{22} \right\|_q \leq 6.5 \|A_{-\max(r)}\|_q T_R \left(\frac{1}{\sigma_{\min}(V_{11})} + 1 \right)$$

with probability at least $1 - e^{-\delta m_1}$.

Parallel results hold for the case when U is fixed and V has orthonormal columns that are randomly sampled based on the Haar measure, and we observe the first m_1 rows and first m_2 columns of A . Assume there exists $r < \frac{1}{2} \min(m_1, m_2)$ such that

$$\sigma_{r+1}(A) \leq \frac{1}{5} \sigma_r(A) \sigma_{\min}(U_{11}) \sqrt{\frac{m_2}{p_2}}.$$

Then there exist uniform constants $c, \delta > 0$ such that if $m_2 \geq cr$, \hat{A}_{22} is given by Algorithm 2 with column thresholding with the break condition $\|D_{C,s}\| \leq T_C$, where $T_C \geq 2\sqrt{\frac{p_2}{m_2}}$, we have for all $1 \leq q \leq \infty$,

$$\left\| \hat{A}_{22} - A_{22} \right\|_q \leq 6.5 \|A_{-\max(r)}\|_q T_C \left(\frac{1}{\sigma_{\min}(U_{11})} + 1 \right)$$

with probability at least $1 - e^{-\delta m_2}$.

3.4 Simulation

In this section, we show results from extensive simulation studies that examine the numerical performance of Algorithm 2 on randomly generated matrices for various values of p_1 , p_2 , m_1 and m_2 . We first consider settings where a gap between some adjacent singular values exists, as required by our theoretical analysis. Then we investigate settings where the singular values decay smoothly with no significant gap between adjacent singular values. The results show that the proposed procedure performs well even when there is no significant gap, as long as the singular values decay at a reasonable rate.

We also examine how sensitive the proposed estimators are to the choice of the threshold and the choice between row and column thresholding. In addition, we com-

pare the performance of the SMC method with that of the NNM method. Finally, we consider a setting similar to the real data application discussed in the next section. Results shown below are based on 200-500 replications for each configuration. Additional simulation results on the effect of m_1 , m_2 and ratio p_1/m_1 are provided in the Appendix. Throughout, we generate the random matrix A from $A = U\Sigma V$, where the singular values of the diagonal matrix Σ are chosen accordingly for different settings. The singular spaces U and V are drawn randomly from the Haar measure. Specifically, we generate i.i.d. standard Gaussian matrix $\tilde{U} \in \mathbb{R}^{p_1 \times \min(p_1, p_2)}$ and $\tilde{V} \in \mathbb{R}^{p_2 \times \min(p_1, p_2)}$, then apply the QR decomposition to \tilde{U} and \tilde{V} and assign U and V with the Q part of the result.

We first consider the performance of Algorithm 2 when a significant gap between the r^{th} and $(r + 1)^{\text{th}}$ singular values of A . We fixed $p_1 = p_2 = 1000$, $m_1 = m_2 = 50$ and choose the singular values as

$$\underbrace{\{1, \dots, 1\}}_r, g^{-1}1^{-1}, g^{-1}2^{-1}, \dots, \quad g = 1, 2, \dots, 10, \quad r = 4, 12 \text{ and } 20. \quad (3.21)$$

Here r is the rank of the major low-rank part $A_{\max(r)}$, $g = \frac{\sigma_r(A)}{\sigma_{r+1}(A)}$ is the gap ratio between the r^{th} and $(r + 1)^{\text{th}}$ singular values of A . The average loss of \hat{A}_{22} from Algorithm 2 with the row thresholding and $T_R = 2\sqrt{p_1/m_1}$ under both the spectral norm and Frobenius norm losses are given in Figure 3.3. The results suggest that our algorithm performs better when r gets smaller and gap ratio $g = \sigma_r(A)/\sigma_{r+1}(A)$ gets larger. Moreover, even when $g = 1$, namely there is no significant gap between any adjacent singular values, our algorithm still works well for small r . As will be seen in the following simulation studies, this is generally the case as long as the singular values of A decay sufficiently fast.

We now turn to the settings with the singular values being $\{j^{-\alpha}\}_{j=1}^{\min(p_1, p_2)}$ and

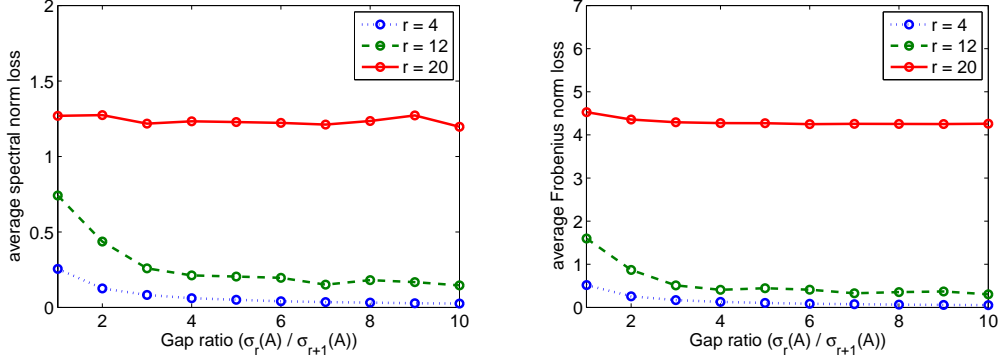


Figure 3.3: Spectral norm loss (left panel) and Frobenius norm loss (right panel) when there is a gap between $\sigma_r(A)$ and $\sigma_{r+1}(A)$. The singular value values of A are given by (3.21), $p_1 = p_2 = 1000$, and $m_1 = m_2 = 50$.

various choices of α , p_1 and p_2 . Hence, no significant gap between adjacent singular values exists under these settings and we aim to demonstrate that our method continues to work well. We first consider $p_1 = p_2 = 1000$, $m_1 = m_2 = 50$ and let α range from 0.3 to 2. Under this setting, we also study how the choice of thresholds affect the performance of our algorithm. For simplicity, we report results only for row thresholding as results for column thresholding are similar. The average loss of \hat{A}_{22} from Algorithm 2 with $T_R \in \{c\sqrt{m_1/p_1}, c \in [1, 6]\}$ under both the spectral norm and Frobenius norm are given in Figure 3.4. In general, the algorithm performs well provided that α is not too small and as expected, the average loss decreases with a higher decay rate in the singular values. This indicates that the existence of a significant gap between adjacent singular values is not necessary in practice, provided that the singular values decay sufficiently fast. When comparing the results across different choices of the threshold, $c = 2$ as suggested in our theoretical analysis is indeed the optimal choice. Thus, in all subsequent numerical analysis, we fix $c = 2$.

To investigate the impact of row versus column thresholding, we let the singular value decay rate be $\alpha = 1$, $p_1 = 300$, $p_2 = 3000$, and m_1 and m_2 varying from 10

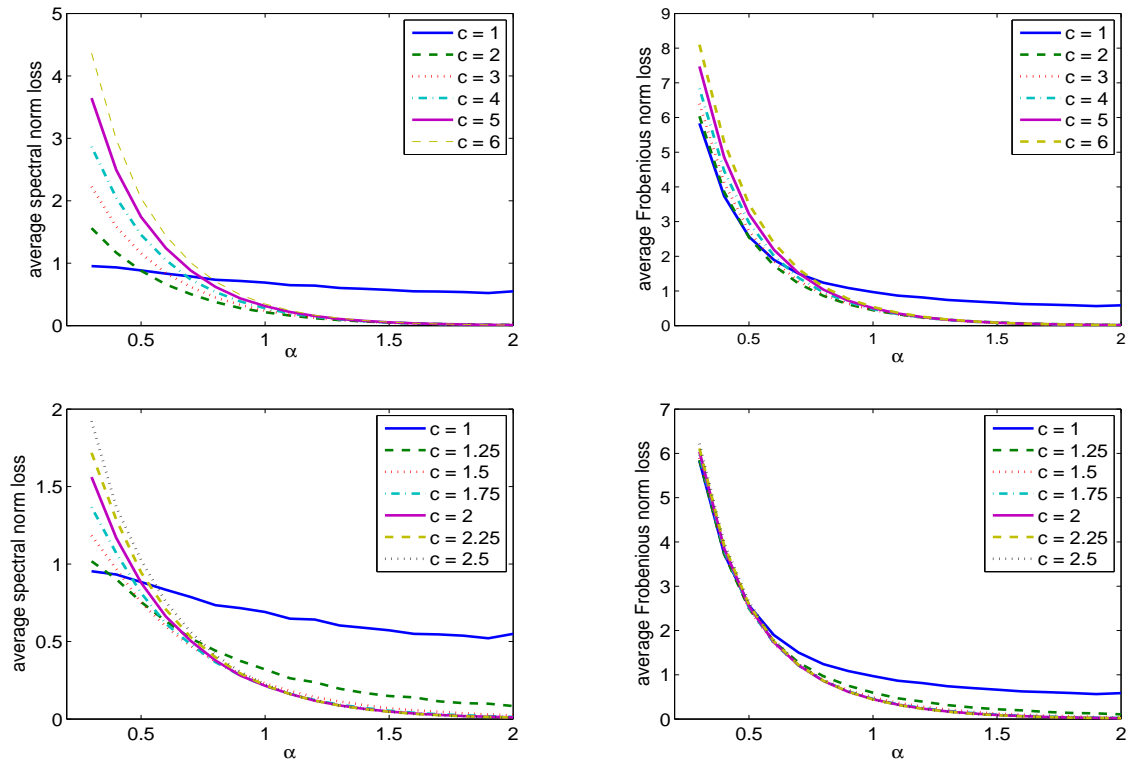


Figure 3.4: Spectral norm loss (left panel) and Frobenius norm loss (right panel) as the thresholding constant c varies. The singular values of A are $\{j^{-\alpha}, j = 1, 2, \dots\}$ with α varying from 0.3 to 2, $p_1 = p_2 = 1000$, and $m_1 = m_2 = 50$.

to 150. The original matrix A is generated the same way as before. We apply row and column thresholding with $T_R = 2\sqrt{p_1/m_1}$ and $T_C = 2\sqrt{p_2/m_2}$. It can be seen from Figure 3.5 that when the observed rows and columns are selected randomly, the results are not sensitive to the choice between row and column thresholding.

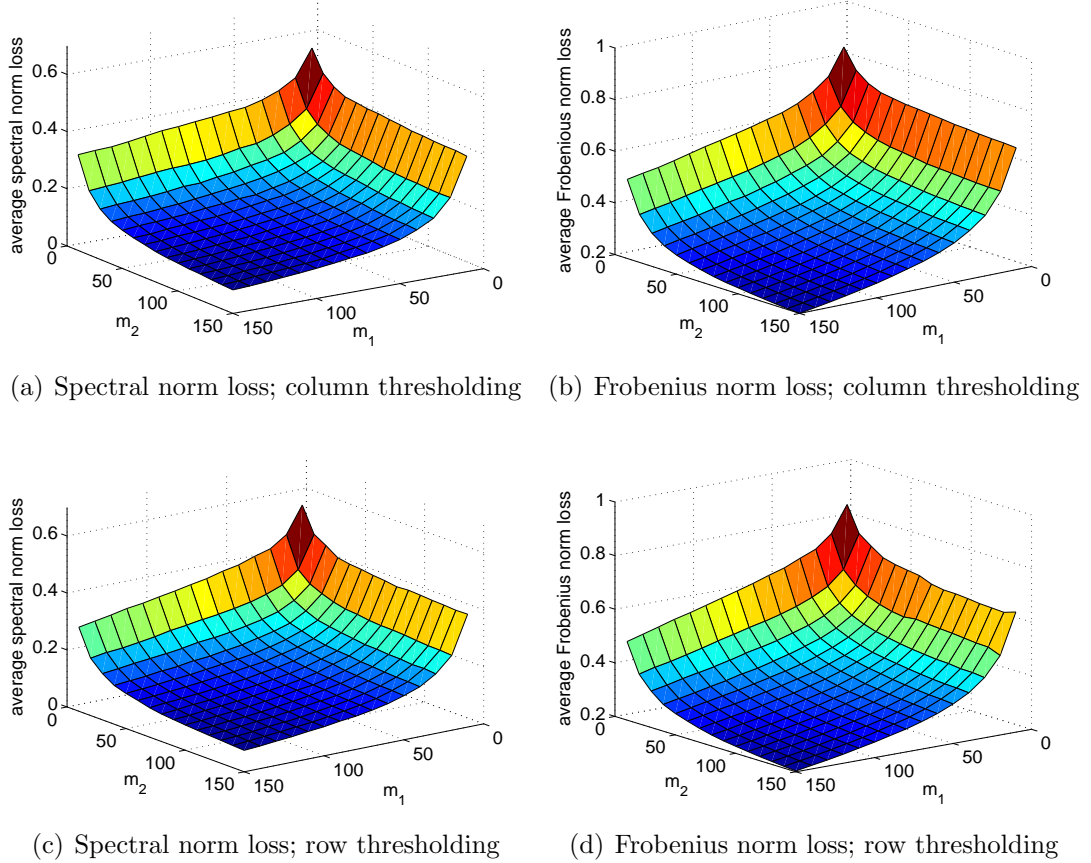


Figure 3.5: Spectral and Frobenius norm losses with column/row thresholding. The singular values of A are $\{j^{-1}, j = 1, 2, \dots\}$, $p_1 = 300$, $p_2 = 3000$, and $m_1, m_2 = 10, \dots, 150$.

We next turn to the comparison between our proposed SMC algorithm and the penalized NNM method which recovers A by (3.4). The solution to (3.4) can be solved by the spectral regularization algorithm by Mazumder et al. (2010) or the accelerated proximal gradient algorithm by Toh and Yun (2010), where these two

methods provide similar results. We use 5-fold cross-validation to select the tuning parameter t . Details on the implementation can be found in the Appendix.

We consider the setting where $p_1 = p_2 = 500$, $m_1 = m_2 = 50, 100$ and the singular value decay rate α ranges from 0.6 to 2. As shown in Figure 3.6, the proposed SMC method substantially outperform the penalized NNM method with respect to both the spectral and Frobenius norm loss, especially as α increases.

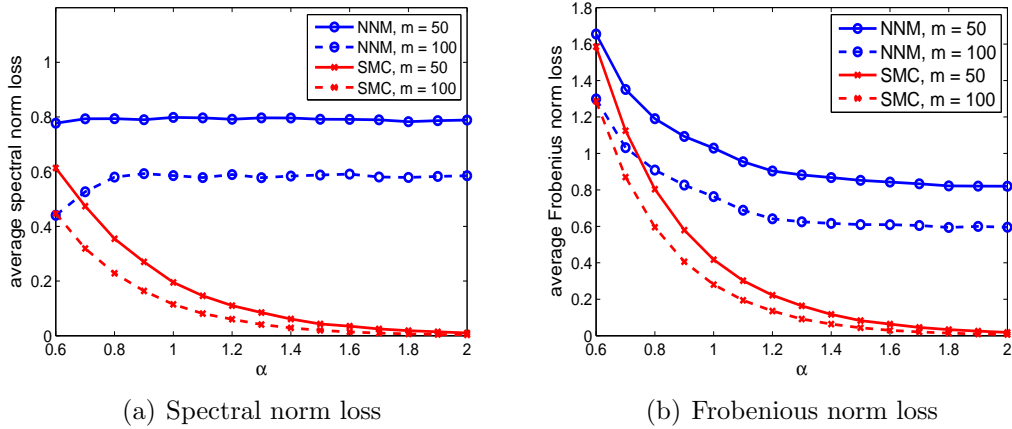


Figure 3.6: Comparison of the proposed SMC method with the NNM method with 5-cross-validation for the settings with singular values of A being $\{j^{-\alpha}, j = 1, 2, \dots\}$ for α ranging from 0.6 to 2, $p_1 = p_2 = 500$, and $m_1 = m_2 = 50$ or 100.

Finally, we consider a simulation setting that mimics the ovarian cancer data application considered in the next section, where $p_1 = 1148$, $p_2 = 1225$, $m_1 = 230$, $m_2 = 426$ and the singular values of A decay at a polynomial rate α . Although the singular values of the full matrix are unknown, we estimate the decay rate based on the singular values of the fully observed 552 rows of the matrix from the TCGA study, denoted by $\{\sigma_j, j = 1, \dots, 522\}$. A simple linear regression of $\{\log(\sigma_j), j = 1, \dots, 522\}$ on $\{\log(j), j = 1, \dots, 522\}$ estimates α as 0.8777. In the simulation, we randomly generate $A \in \mathbb{R}^{p_1 \times p_2}$ such that the singular values are fixed as $\{j^{-.8777}, j = 1, 2, \dots\}$. For comparison, we also obtained results for $\alpha = 1$ as well as those based

on the penalized NNM method with 5-cross-validation. As shown in Table 3.1, the relative spectral norm loss and relative Frobenius norm loss of the proposed method are reasonably small and substantially smaller than those from the penalized NNM method.

	Relative spectral norm loss		Relative Frobenius norm loss	
	SMC	NNM	SMC	NNM
$\alpha = 0.8777$	0.1253	0.4614	0.2879	0.6122
$\alpha = 1$	0.0732	0.4543	0.1794	0.5671

Table 3.1: Relative spectral norm loss ($\|\hat{A}_{22} - A_{22}\|/\|A_{22}\|$) and Frobenius norm loss ($\|\hat{A}_{22} - A_{22}\|_F/\|A_{22}\|_F$) for $p_1 = 1148$, $p_2 = 1225$, $m_1 = 230$, $m_2 = 426$ and singular values of A being $\{j^{-\alpha} : j = 1, 2, \dots\}$.

3.5 Application in Genomic Data Integration

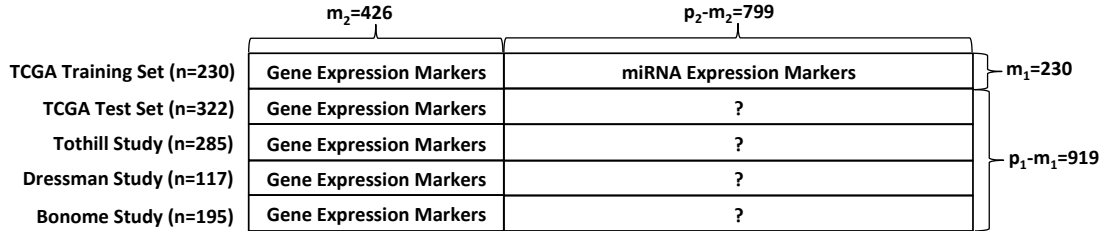
In this section, we apply our proposed procedures to integrate multiple genomic studies of ovarian cancer (OC). OC is the fifth leading cause of cancer mortality among women, attributing to 14,000 deaths annually (Siegel et al., 2013). OC is a relatively heterogeneous disease with 5-year survival rate varying substantially among different subgroups. The overall 5-year survival rate is near 90% for stage I cancer. But the majority of the OC patients are diagnosed as stage III/IV diseases and tend to develop resistance to chemotherapy, resulting a 5-year survival rate only about 30% (Holschneider and Berek, 2000). On the other hand, a small minority of advanced cancers are sensitive to chemotherapy and do not relapse after treatment completion. Such a heterogeneity in disease progression is likely to be in part attributable to variations in underlying biological characteristics of OC (Berchuck et al., 2005). This heterogeneity and the lack of successful treatment strategies motivated multi-

ple genomic studies of OC to identify molecular signatures that can distinguish OC subtypes, and in turn help to optimize and personalize treatment. For example, the Cancer Genome Atlas (TCGA) comprehensively measured genomic and epigenetic abnormalities on high grade OC samples (Cancer Genome Atlas Research Network, 2011). A gene expression risk score based on 193 genes, \mathcal{G} , was trained on 230 training samples, denoted by $\text{TCGA}^{(t)}$, and shown as highly predictive of OC survival when validated on the TCGA independent validation set of size 322, denoted by $\text{TCGA}^{(v)}$, as well as on several independent OC gene expression studies including those from Bonome et al. (2005) (BONO), Dressman et al. (2007) (DRES) and Tothill et al. (2008) (TOTH).

The TCGA study also showed that clustering of miRNA levels overlaps with gene-expression based clusters and is predictive of survival. It would be interesting to examine whether combining miRNA with \mathcal{G} could improve survival prediction when compared to \mathcal{G} alone. One may use $\text{TCGA}^{(v)}$ to evaluate the added value of miRNA. However, $\text{TCGA}^{(v)}$ is of limited sample size. Furthermore, since miRNA was only measured for the TCGA study, its utility in prediction cannot be directly validated using these independent studies. Here, we apply our proposed SMC method to impute the missing miRNA values and subsequently construct prediction rules based on both \mathcal{G} and the imputed miRNA, denoted by $\widehat{\text{miRNA}}$, for these independent validation sets. To facilitate the comparison with the analysis based on $\text{TCGA}^{(v)}$ alone where miRNA measurements are observed, we only used the miRNA from $\text{TCGA}^{(t)}$ for imputation and reserved the miRNA data from $\text{TCGA}^{(v)}$ for validation purposes. To improve the imputation, we also included additional 300 genes that were previously used in a prognostic gene expression signature for predicting ovarian cancer survival (Denkert et al., 2009). This results in a total of $m_1 = 426$ unique gene expression variables available for imputation. Detailed information on the data used for imputation is

shown in Figure 3.7. Prior to imputation, all gene expression and miRNA levels are log transformed and centered to have mean zero within each study to remove potential platform or batch effects. Since the observable rows (indexing subjects) can be viewed as random whereas the observable columns (indexing genes and miRNAs) are not random, we used row thresholding with threshold $T_R = 2\sqrt{p_1/m_1}$ as suggested in the theoretical and simulation results. For comparison, we also imputed data using the penalized NNM method with tuning parameter t selected via 5-fold cross-validation.

Figure 3.7: Imputation scheme for integrating multiple OC genomic studies.



We first compared $\widehat{\text{miRNA}}$ to the observed miRNA on TCGA^(v). Our imputation yielded a rank 2 matrix for $\widehat{\text{miRNA}}$ and the correlations between the two right and left singular vectors $\widehat{\text{miRNA}}$ to that of the observed miRNA variables are .90, .71, .34, .14, substantially higher than that of those from the NNM method, with the corresponding values 0.45, 0.06, 0.10, 0.05. This suggests that the SMC imputation does a good job in recovering the leading projections of the miRNA measurements and outperforms the NNM method.

To evaluate the utility of $\widehat{\text{miRNA}}$ for predicting OC survival, we used the TCGA^(t) to select 117 miRNA markers that are marginally associated with survival with a nominal p -value threshold of .05. We use the two leading principal components (PCs) of the 117 miRNA markers, $\text{miRNA}^{\text{PC}} = (\text{miRNA}_1^{\text{PC}}, \text{miRNA}_2^{\text{PC}})^T$, as predictors for the survival outcome in addition to \mathcal{G} . The imputation enables us to integrate information from 4 studies including TCGA^(t), which could substantially improve efficiency and

prediction performance. We first assessed the association between $\{\text{miRNA}^{\text{PC}}, \mathcal{G}\}$ and OC survival by fitting a stratified Cox model (Kalbfleisch and Prentice, 2011) to the integrated data that combines $\text{TCGA}^{(v)}$ and the three additional studies via either the SMC or NNM methods. In addition, we fit the Cox model to (i) $\text{TCGA}^{(v)}$ set alone with miRNA^{PC} obtained from the observed miRNA; and (ii) each individual study separately with imputed miRNA^{PC} . As shown in Table 3.2(a), the log hazard ratio (logHR) estimates for miRNA^{PC} from the integrated analysis, based on both SMC and NNM methods, are similar in magnitude to those obtained based on the observed miRNA values with $\text{TCGA}^{(v)}$. However, the integrated analysis has substantially smaller standard error (SE) estimates due to the increased sample sizes. The estimated logHRs are also reasonably consistent across studies when separate models were fit to individual studies.

We also compared the prediction performance of the model based on \mathcal{G} alone to the model that includes both \mathcal{G} and the imputed miRNA^{PC} . Combining information from all 4 studies via standard meta analysis, the average improvement in C-statistic was 0.032 (SE = 0.013) for the SMC method and 0.001 (SE = 0.009) for the NNM method, suggesting that the imputed miRNA^{PC} from the SMC method has much higher predictive value compared to those obtained from the NNM method.

In summary, the results shown above suggest that our SMC procedure accurately recovers the leading PCs of the miRNA variables. In addition, adding miRNA^{PC} obtained from imputation using the proposed SMC method could significantly improve the prediction performance, which confirms the value of our method for integrative genomic analysis. When comparing to the NNM method, the proposed SMC method produces summaries of miRNA that is more correlated with the truth and yields leading PCs that are more predictive of OC survival.

Table 3.2: Shown in (a) are the estimates of the log hazard ratio (logHR) along with their corresponding standard errors (SE) and p -values by fitting stratified Cox model integrating information from 4 independent studies with imputed miRNA based on the SMC method and the nuclear norm minimization (NNM); and Cox model to the TCGA test data with original observed miRNA (Ori.). Shown also are the estimates for each individual studies by fitting separate Cox models with imputed miRNA.

(a) Integrated Analysis with Imputed miRNA vs Single study with observed miRNA

	logHR			SE			p -value		
	Ori.	SMC	NNM	Ori.	SMC	NNM	Ori.	SMC	NNM
\mathcal{G}	.067	.143	.168	.041	.034	.028	.104	.000	.000
miRNA ₁ ^{PC}	-.012	-.019	-.013	.009	.006	.012	.218	.001	.283
miRNA ₂ ^{PC}	.023	.018	-.005	.014	.009	.014	.092	.039	.725

(b) Estimates for Individual Studies with Imputed miRNA from the SMC method

	logHR				SE				p -value			
	TCGA	TOTH	DRES	BONO	TCGA	TOTH	DRES	BONO	TCGA	TOTH	DRES	BONO
\mathcal{G}	.051	.377	.174	.311	.048	.069	.132	.117	.286	.000	.187	.008
miRNA ₁ ^{PC}	-.014	-.021	-.031	-.010	.011	.012	.014	.014	.207	.082	.030	.484
miRNA ₂ ^{PC}	.014	.045	-.021	.036	.016	.018	.022	.019	.391	.009	.336	.054

(c) Estimates for Individual Studies with Imputed miRNA from the NNM method

	logHR				SE				p -value			
	TCGA	TOTH	DRES	BONO	TCGA	TOTH	DRES	BONO	TCGA	TOTH	DRES	BONO
\mathcal{G}	.082	.405	.361	.258	.037	.066	.114	.088	.028	.000	.002	.003
miRNA ₁ ^{PC}	-.045	.016	.055	-.008	.021	.026	.031	.023	.034	.544	.076	.721
miRNA ₂ ^{PC}	.008	-.086	-.043	.019	.026	.027	.034	.029	.758	.002	.201	.496

3.6 Discussions

The present chapter introduced a new framework of SMC where a subset of the rows and columns of an approximately low-rank matrix are observed. We proposed an SMC method for the recovery of the whole matrix with theoretical guarantees. The proposed procedure significantly outperforms the conventional NNM method for matrix completion, which does not take into account the special structure of the observations. As shown by our theoretical and numerical analyses, the widely adopted NNM methods for matrix completion are not suitable for the SMC setting. These NNM methods perform particularly poorly when a small number of rows and columns are observed.

The key assumption in matrix completion is the matrix being approximately low rank. This is reasonable in the ovarian cancer application since as indicated in the results from the TCGA study (Cancer Genome Atlas Research Network, 2011), the patterns observed in the miRNA signature are highly correlated with the patterns observed in the gene expression signature. This suggests the high correlation among the selected gene expression and miRNA variables. Results from the imputation based on the approximate low rank assumption given in Section 3.5 are also encouraging with promising correlations with true signals and good prediction performance from the imputed miRNA signatures. We expect that this imputation method will also work well in genotyping and sequencing applications, particularly for regions with reasonably high linkage disequilibrium.

Another main assumption that is needed in the theoretical analysis is that there is a significant gap between the r^{th} and $(r + 1)^{th}$ singular values of A . This assumption may not be valid in real practice. In particular, the singular values of the ovarian dataset analyzed in Section 3.5 is decreasing smoothly without a significant gap.

However, it has been shown in the simulation studies presented in Section 3.4 that, although there is no significant gap between any adjacent singular values of the matrix to be recovered, the proposed SMC method works well as long as the singular values decay sufficiently fast. Theoretical analysis for the proposed SMC method under more general patterns of singular value decay warrants future research.

To implement the proposed Algorithm 2, major decisions include the choice of threshold values and choosing between column thresholding and row thresholding. Based on both theoretical and numerical studies, optimal threshold values can be set as $T_C = 2\sqrt{p_2/m_2}$ for column thresholding and $T_R = 2\sqrt{p_1/m_1}$ for row thresholding. Simulation results in Section 3.4 show that when both rows and columns are randomly chosen, the results are very similar. In the real data applications, the choice between row thresholding and column thresholding depends on whether the rows or columns are more “homogeneous”, or closer to being randomly sampled. For example, in the ovarian cancer dataset analyzed in Section 3.5, the rows correspond to the patients and the columns correspond to the gene expression levels and miRNA levels. Thus the rows are closer to random sample than the columns, consequently it is more natural to use the row thresholding in this case.

We have shown both theoretically and numerically in Sections 3.3 and 3.4 that Algorithm 2 provides a good recovery of A_{22} . However, the naive implementation of this algorithm requires $\min(m_1, m_2)$ matrix inversions and multiplication operations in the for loop that calculates $\|D_{R,s}\|$ (or $\|D_{C,s}\|$), $s \in \{\hat{r}, \hat{r} + 1, \dots, \min(m_1, m_2)\}$. Taking into account the relationship among $D_{R,s}$ (or $D_{C,s}$) for different s 's, it is possible to simultaneously calculate all $\|D_{R,s}\|$ (or $\|D_{C,s}\|$) and accelerate the computations. For reasons of space, we leave optimal implementation of Algorithm 2 as future work.

My future research will be continuation of my current work. Beyond that, I also have some newer problems in mind. I am hoping to expand my statistical interests and get into more collaborative work with people in different fields. The following are some projects that I am interested in working on in the future.

4.1 Inference for Large Gaussian Graphical Model with Missing Data

The Gaussian graphical model is a powerful tool in modern statistics for analyzing relationship networks. In the era of high-dimensional statistics, missing data also occurs so that the traditional statistical inference methods often no longer apply. Therefore, it would be very interesting if one can develop a set of methods which make statistical inference on large Gaussian graphical model with missing data.

Within this direction, so far we have preliminary results on estimating a large sparse precision matrix with incomplete data by constraint ℓ_1 -norm minimization. However, more fundamental problems in this area still remain unclear, such as how to recover the support or make inference on the Gaussian graph. I think such problems are important and I shall make future efforts.

4.2 High-dimensional Sparsity Test

In modern high-dimensional statistics, different kinds of structural assumptions have been imposed on the model. Among those assumptions, sparsity of the object is one of the most widely used and becomes the foundation of many methodologies and theories. However, it is sometimes unclear whether this assumption is valid or not. For example, gene transcription networks often contain the so-called “hub nodes” where the corresponding gene expressions are correlated with many other gene expressions, and may thus yield a non-sparse network. It would be a very interesting result if we can develop a set of methodologies to test if the object is sparse or not. The object here can range from linear regression coefficient to the covariance structure or network, e.g. detection of the hub node. Up to now we have preliminary results in high-dimensional regression showing that a certain Chi-squared test is powerful for a sparse null against a non-sparse alternative. I would like to make much more efforts towards solving this problem.

4.3 Noisy Structured Matrix Completion

Same as the structured matrix completion setting (Chapter 3), when the observations are with i.i.d. noise, the original algorithm proposed in Chapter 3 may be sub-optimal. The main reason is that, by random matrix theory, the i.i.d. noise perturbation is proportionally spread into each block with high probability and a better denoising rule should be applied. A possible path to improvement is to combine matrix denoising via SVD with our proposed SMC. The analysis of the detailed algorithm is an ongoing work.

A

Appendices

A.1 Supplement for Chapter 1

We shall prove the main results for Chapter 1 in this Appendix section.

A.1.1 Proof of Lemma 1.1.1.

First, suppose $v \in T(\alpha, s)$. We can prove v is in the convex hull of $U(\alpha, s, v)$ by induction. If v is s -sparse, v itself is in $U(\alpha, s, v)$.

Suppose the statement is true for all $(l-1)$ -sparse vectors v ($l-1 \geq s$). Then for any l -sparse vector v such that $\|v\|_\infty \leq \alpha$, $\|v\|_1 \leq s\alpha$, without loss of generality we assume that v is not $(l-1)$ -sparse (otherwise the result holds by assumption of $l-1$). Hence we can express v as $v = \sum_{i=1}^l a_i e_i$, where e_i 's are different unit vectors with one entry of ± 1 and other entries of zeros; $a_1 \geq a_2 \geq \dots \geq a_l > 0$. Since $\sum_{i=1}^l a_i = \|v\|_1 \leq s\alpha$, so

$$1 \in D \triangleq \{1 \leq j \leq l-1 : a_j + a_{j+1} + \dots + a_l \leq (l-j)\alpha\},$$

which means D is not empty. Take the largest element in D as j , which implies

$$\begin{aligned} a_j + a_{j+1} + \dots + a_l &\leq (l-j)\alpha, \\ a_{j+1} + a_{j+2} + \dots + a_l &> (l-j-1)\alpha. \end{aligned} \tag{A.1}$$

(It is noteworthy that even if the largest j in D is $l-1$, (A.1) still holds). Define

$$b_w \triangleq \frac{\sum_{i=j}^l a_i}{l-j} - a_w, \quad j \leq w \leq l, \tag{A.2}$$

which satisfies $\sum_{i=j}^l a_i = (l-j) \sum_{i=j}^l b_i$. By (A.1), for all $j \leq w \leq l$,

$$b_w \geq b_j = \frac{\sum_{i=j+1}^l a_i}{l-j} - \frac{l-j-1}{l-j} a_j \geq \frac{\sum_{i=j+1}^l a_i - (l-j-1)\alpha}{l-j} > 0.$$

In addition, we define

$$\begin{aligned} v_w &\triangleq \sum_{i=1}^{j-1} a_i e_i + \left(\sum_{i=j}^l b_i \right) \sum_{i=j, i \neq w}^l e_i \in \mathbb{R}^p, \\ \lambda_w &\triangleq \frac{b_w}{\sum_{i=j}^l b_i}, \quad j \leq w \leq l, \end{aligned} \tag{A.3}$$

then $0 \leq \lambda_w \leq 1$, $\sum_{w=j}^l \lambda_w = 1$, $\sum_{w=j}^l \lambda_w v_w = v$, $\text{supp}(v_w) \subseteq \text{supp}(v)$. We also have

$$\|v_w\|_1 = \sum_{i=1}^{j-1} a_i + (l-j) \sum_{w=j}^l b_w = \sum_{i=1}^{j-1} a_i + \sum_{i=j}^l a_i = \|v\|_1,$$

$$\|v_w\|_\infty = \max\{a_1, \dots, a_{j-1}, \sum_{i=j}^l b_i\} \leq \max\left\{\alpha, \frac{\sum_{i=j}^l a_i}{l-j}\right\} \leq \alpha.$$

The last inequality is due to the first part of (A.1). Finally, note that v_w is $(l-1)$ -sparse, we can use the induction assumption to find $\{u_{i,w} \in \mathbb{R}^p, \lambda_{i,w} \in \mathbb{R} : 1 \leq i \leq N_w, j \leq w \leq l\}$ such that

$$u_{i,w} \text{ is } s\text{-sparse, } \text{supp}(u_{i,w}) \subseteq \text{supp}(v_i) \subseteq \text{supp}(v),$$

$$\|u_{i,w}\|_1 = \|v_i\|_1 = \|v\|_1, \quad \|u_{i,w}\|_\infty \leq \alpha;$$

In addition, $v_i = \sum_{i=1}^{N_w} \lambda_{i,w} u_{i,w}$, so $v = \sum_{w=j}^l \sum_{i=1}^{N_w} \lambda_w \lambda_{i,w} u_{i,w}$, which proves the result for l .

The proof of the other part of the lemma is easier. When v is in the convex hull

of $U(\alpha, s, v)$, then we have

$$\|v\|_\infty = \left\| \sum_{i=1}^N \lambda_i u_i \right\|_\infty \leq \sum_{i=1}^N \lambda_i \|u_i\|_\infty \leq \alpha,$$

$$\|v\|_1 = \left\| \sum_{i=1}^N \lambda_i u_i \right\|_1 \leq \sum_{i=1}^N \lambda_i \|u_i\|_1 \leq \sum_{i=1}^N \lambda_i \|u_i\|_0 \|u_i\|_\infty \leq s\alpha,$$

which finished the proof of the lemma. \square

A.1.2 Proof of Theorems 1.1.1 and 1.2.2

We first state two lemmas. One important technical tool we will use is the following Division Lemma. In order to relate the general vectors and matrices with the RIP condition whose constraint is on the sparse vectors and low-rank matrices, a natural approach is to divide these elements into sums of sparse or low-rank components. Consequently, we introduce the Division Lemma below, which is a key technical tool for the proof of sufficiency of $\delta_k^A < 1/3$ and $\delta_r^M < 1/3$.

Lemma A.1.1 (Division Lemma). *Let r and m be positive integers with $m \geq 2r$. Let $a_1 \geq a_2 \geq a_3 \geq \dots \geq a_m \geq 0$ be a sequence of non-increasing real numbers satisfying*

$$\sum_{w=1}^r a_w \geq \sum_{w=r+1}^m a_w. \quad (\text{A.4})$$

Then there exist non-negative real numbers $\{s_{ij}\}_{1 \leq i \leq r, 2r+1 \leq j \leq m}$ such that

$$\sum_{i=1}^r s_{ij} = a_j, \quad \forall 2r+1 \leq j \leq m, \quad (\text{A.5})$$

and

$$\frac{1}{r} \sum_{w=1}^r a_w \geq a_{r+i} + \sum_{j=2r+1}^m s_{ij}, \quad \forall 1 \leq i \leq r. \quad (\text{A.6})$$

The proof of Lemma A.1.1 is simply by induction on m . The Division Lemma can be illustrated as in the following table. Each row is an inequality; every element in the first row equals the sum of remaining elements in the same column:

a_1	a_2	\cdots	a_r	\geq	a_{r+1}	a_{r+2}	\cdots	a_{2r}	+	a_{2r+1}	\cdots	a_m
a_1/r	a_2/r	\cdots	a_r/r	\geq	a_{r+1}				+	$s_{1,2r+1}$	\cdots	$s_{1,m}$
a_1/r	a_2/r	\cdots	a_r/r	\geq		a_{r+2}			+	$s_{2,2r+1}$	\cdots	$s_{2,m}$
\vdots	\vdots	\ddots	\vdots	\geq			\ddots		+	\vdots		\vdots
a_1/r	a_2/r	\cdots	a_r/r	\geq				a_{2r}	+	$s_{r,2r+1}$	\cdots	$s_{r,m}$

Lemma A.1.2, which characterizes the null space properties, is from Stojnic et al. (2008) and Oymak and Hassibi (2010).

Lemma A.1.2. *In the noiseless case, using (1.2) with $\mathcal{B} = \{0\}$ one can recover all k -sparse signals β if and only if for all $h \in \mathcal{N}(A) \setminus \{0\}$,*

$$2\|h_{\max(k)}\|_1 < \|h\|_1.$$

Similarly in the noiseless case, using (1.4) with $\mathcal{B} = \{0\}$ one can recover all matrices X of rank at most r if and only if for all $R \in \mathcal{N}(\mathcal{M}) \setminus \{0\}$,

$$2\|R_{\max(r)}\|_* < \|R\|_*.$$

The key to the proof of this theorem is parallelogram identity, since it provides equality rather than inequality in the estimation in ℓ_2 norm as we shall see later. The proof of Theorems 1.1.1 and 1.2.2 shall be divided into three parts: $\delta_k^A < 1/3$, $\delta_k^A + \theta_{k,k}^A < 1$ (or $\delta_a^A + C_{a,b,k}\theta_{a,b}^A < 1$) and $\delta_{tk}^A < \sqrt{(t-1)/t}$.

Part 1. If $\delta_k^A < 1/3$ for some $k \geq 2$.

By Lemma A.1.2, we only need to show for all $\beta \in \mathcal{N}(A) \setminus \{0\}$, it satisfies $\|\beta_{\max(k)}\|_1 < \|\beta_{-\max(k)}\|_1$.

For the convenience of presentation, we call a vector with 1 or -1 in only one entry and zeros elsewhere as the *indicator vector*.

Suppose there exists $h \in \mathcal{N}(A) \setminus \{0\}$ such that $\|h_{\max(k)}\|_1 < \|h_{-\max(k)}\|_1$. Then h can be written as

$$h = \sum_{i=1}^p a_i u_i$$

where $\{u_i\}_{i=1}^p$ are indicator vectors with different support in \mathbb{R}^p ; $\{a_i\}_{i=1}^p$ is a non-negative and decreasing sequence. Since we can set $a_i = 0$ if $i \geq p$, without loss of generality we can assume that $p \geq k$.

By Lemma A.1.1, we can find $\{s_{ij}\}_{1 \leq i \leq k, 2k+1 \leq j \leq p}$ satisfying (A.5) and (A.6) with a modification of notations.

1. When k is even, suppose

$$\begin{aligned} h_{11} &= \sum_{i=1}^{k/2} a_i u_i, & h_{12} &= \sum_{i=k/2+1}^k a_i u_i, & h_{21} &= \sum_{i=k+1}^{3k/2} a_i u_i, & h_{22} &= \sum_{i=3k/2+1}^{2k} a_i u_i \\ h_{31} &= \sum_{j=2k+1}^p \left(\sum_{i=1}^{k/2} s_{ij} u_j \right), & h_{32} &= \sum_{j=2k+1}^p \left(\sum_{i=k/2+1}^k s_{ij} u_j \right) \end{aligned} \tag{A.7}$$

then $A(h_{11} + h_{12} + h_{21} + h_{22} + h_{31} + h_{32}) = Ah = 0$. By the parallelogram

identity,

$$\begin{aligned}
& \|A(-h_{11} + h_{22} + h_{32})\|^2 + \|A(-h_{12} + h_{21} + h_{31})\|^2 \\
&= \frac{1}{2} \|A(-h_{11} - h_{12} + h_{21} + h_{22} + h_{31} + h_{32})\|^2 \\
&\quad + \frac{1}{2} \|A(-h_{11} + h_{12} - h_{21} + h_{22} - h_{31} + h_{32})\|^2 \\
&= \frac{1}{2} \|A(2h_{11} + 2h_{12})\|^2 + \frac{1}{4} \|A(-2h_{11} - 2h_{21} - 2h_{31})\|^2 \\
&\quad + \frac{1}{4} \|A(2h_{12} + 2h_{22} + 2h_{32})\|^2 \\
&= 2\|A(h_{11} + h_{12})\|^2 + \|A(h_{11} + h_{21} + h_{31})\|^2 + \|A(h_{12} + h_{22} + h_{32})\|^2
\end{aligned} \tag{A.8}$$

Similarly as the matrix case, we use Lemma A.1.4 and get

$$\begin{aligned}
& \|A(h_{11} + h_{21} + h_{31})\|^2 - \|A(-h_{12} + h_{21} + h_{31})\|^2 \\
&\geq (1 - \delta_k^A) \left(\sum_{i=1}^{k/2} a_i^2 + \sum_{i=k+1}^{3k/2} (a_i + \sum_{j=2k+1}^p s_{ij})^2 \right) \\
&\quad - (1 + \delta_k^A) \left(\sum_{i=k/2+1}^k a_i^2 + \sum_{i=k+1}^{3k/2} (a_i + \sum_{j=2k+1}^p s_{ij})^2 \right)
\end{aligned} \tag{A.9}$$

Similarly,

$$\begin{aligned}
& \|A(h_{12} + h_{22} + h_{32})\|^2 - \|A(-h_{11} + h_{22} + h_{32})\|^2 \\
&\geq (1 - \delta_k^A) \left(\sum_{i=k/2+1}^k a_i^2 + \sum_{i=3k/2+1}^{2k} (a_i + \sum_{j=2k+1}^p s_{ij})^2 \right) \\
&\quad - (1 + \delta_k^A) \left(\sum_{i=1}^{k/2} a_i^2 + \sum_{i=3k/2+1}^{2k} (a_i + \sum_{j=2k+1}^p s_{ij})^2 \right)
\end{aligned} \tag{A.10}$$

Let the right hand side of (A.8) minus the left hand side. Along with (A.9),

(A.10), we get

$$\begin{aligned}
0 &= RHS - LHS \\
&\geq 2(1 - \delta_k^A) \left(\sum_{i=1}^k a_i^2 \right) - 2\delta_k^A \sum_{i=1}^k a_i^2 - 2\delta_k^A \left(\sum_{i=k+1}^{2k} (a_i + \sum_{j=2k+1}^p s_{ij})^2 \right) \\
&\geq 2(1 - 2\delta_k^A) \sum_{i=1}^k a_i^2 - 2\delta_k^A k \left(\frac{\sum_{i=1}^k a_i}{k} \right)^2 \\
&\geq 2(1 - 3\delta_k^A) \sum_{i=1}^k a_i^2
\end{aligned}$$

The last two inequalities are due to (A.6) and Cauchy-Schwarz inequality. It contradicts the fact that $h \neq 0$ and $\delta_k^A < 1/3$.

2. When k is odd, $k \geq 3$, note

$$\begin{aligned}
h_{11} &= a_1 u_1, & h_{12} &= \sum_{i=2}^{(k+1)/2} a_i u_i, & h_{13} &= \sum_{i=(k+3)/2}^k a_i u_i \\
h_{21} &= a_{k+1} u_{k+1}, & h_{22} &= \sum_{i=k+2}^{(3k+1)/2} a_i u_i, & h_{23} &= \sum_{i=(3k+3)/2}^{2k} a_i u_i \\
h_{31} &= \sum_{j=2k+1}^p s_{1j} u_j, & h_{32} &= \sum_{j=2k+1}^p \left(\sum_{i=2}^{(k+1)/2} s_{ij} \right) u_j, & h_{33} &= \sum_{j=2k+1}^p \left(\sum_{i=(k+3)/2}^{2k} s_{ij} \right) u_j
\end{aligned} \tag{A.11}$$

Note $\gamma_1 = -h_{11} + h_{21} + h_{31}$, $\gamma_2 = -h_{12} + h_{22} + h_{23}$, $\gamma_3 = -h_{13} + h_{23} + h_{33}$, we can easily show the following equality

$$\begin{aligned}
&4\|A\gamma_1\|^2 + 4\|A\gamma_2\|^2 + 4\|A\gamma_3\|^2 \\
&= \|A(\gamma_1 + \gamma_2 - \gamma_3)\|^2 + \|A(-\gamma_1 + \gamma_2 + \gamma_3)\|^2 \\
&\quad + \|A(\gamma_1 - \gamma_2 + \gamma_3)\|^2 + \|A(\gamma_1 + \gamma_2 + \gamma_3)\|^2
\end{aligned} \tag{A.12}$$

By the fact that $Ah = 0$, (A.12) means

$$\begin{aligned}
& \|A(-h_{11} + h_{21} + h_{31})\|^2 + \|A(-h_{12} + h_{22} + h_{32})\|^2 + \|A(-h_{13} + h_{23} + h_{33})\|^2 \\
& = \|A(h_{12} + h_{13} + h_{21} + h_{31})\|^2 + \|A(h_{11} + h_{13} + h_{22} + h_{32})\|^2 \\
& \quad + \|A(h_{11} + h_{12} + h_{23} + h_{33})\|^2 + \|A(h_{11} + h_{12} + h_{13})\|^2
\end{aligned} \tag{A.13}$$

Similarly as the even case, by Lemma A.1.4 we have

$$\begin{aligned}
& \|A(h_{12} + h_{13} + h_{21} + h_{31})\|^2 - \|A(-h_{11} + h_{21} + h_{31})\|^2 \\
& \geq (1 - \delta_k^A) \left[\sum_{i=2}^k a_i^2 + (a_{k+1} + \sum_{j=2k+1}^p s_{1,j})^2 \right] - (1 + \delta_k^A) \left[a_1^2 + (a_{k+1} + \sum_{j=2k+1}^p s_{1,j})^2 \right]
\end{aligned} \tag{A.14}$$

$$\begin{aligned}
& \|A(h_{11} + h_{13} + h_{22} + h_{32})\|^2 - \|A(-h_{12} + h_{22} + h_{32})\|^2 \\
& \geq (1 - \delta_k^A) \left[a_1^2 + \sum_{i=(k+3)/2}^k a_i^2 + \sum_{i=2}^{(k+1)/2} (a_i + \sum_{j=2k+1}^p s_{ij})^2 \right] \\
& \quad - (1 + \delta_k^A) \left[\sum_{i=2}^{(k+1)/2} a_i^2 + \sum_{i=2}^{(k+1)/2} (a_i + \sum_{j=2k+1}^p s_{ij})^2 \right]
\end{aligned} \tag{A.15}$$

$$\begin{aligned}
& \|A(h_{11} + h_{12} + h_{23} + h_{33})\|^2 - \|A(-h_{13} + h_{23} + h_{33})\|^2 \\
& \geq (1 - \delta_k^A) \left[\sum_{i=1}^{(k+1)/2} a_i^2 + \sum_{i=(k+3)/2}^k (a_i + \sum_{j=2k+1}^p s_{ij})^2 \right] \\
& \quad - (1 + \delta_k^A) \left[\sum_{i=(k+3)/2}^k a_i^2 + \sum_{i=(k+3)/2}^k (a_i + \sum_{j=2k+1}^p s_{ij})^2 \right]
\end{aligned} \tag{A.16}$$

Let the right hand side of (A.13) minus the left hand side, we can get

$$\begin{aligned}
0 &\geq (1 - \delta_k^A) \left[3 \sum_{i=1}^k a_i^2 + \sum_{i=1}^k (a_{k+i} + \sum_{j=2k+1}^p s_{ij})^2 \right] \\
&\quad - (1 + \delta_k^A) \left[\sum_{i=1}^k a_i^2 + \sum_{i=1}^k (a_{k+i} + \sum_{j=2k+1}^p s_{ij})^2 \right] \\
&= 2 \left[(1 - 2\delta_k^A) \sum_{i=1}^k a_i^2 - \delta_k^A \sum_{i=1}^k (a_{k+i} + \sum_{j=2k+1}^p s_{ij})^2 \right] \\
&\geq 2(1 - 2\delta_k^A) \sum_{i=1}^k a_i^2 - 2\delta_k^A k \left(\frac{\sum_{i=1}^k a_i}{k} \right)^2 \\
&\geq 2(1 - 3\delta_k^A) \sum_{i=1}^k a_i^2
\end{aligned}$$

The last two inequalities are due to (A.6) and Cauchy Schwarz inequality. It contradicts the fact that $h \neq 0$ and $\delta_k^A < 1/3$.

Part 2. If $\delta_k^A + \theta_{k,k}^A < 1$ for some $k \geq 1$. In this scenario, it suffices to prove Theorem 1.2.2 as Theorem 1.1.1 is a special case of Theorem 1.2.2.

As mentioned before, by Oymak et al. (2011), the results for the sparse signal recovery imply the corresponding results for the low-rank matrix recovery. So we will only prove the signal case. By Lemma A.1.2, it suffices to show that for all vectors $h \in \mathcal{N}(A) \setminus \{0\}$, $\|h_{\max(k)}\|_1 < \|h_{-\max(k)}\|_1$.

Suppose there exists $h \in \mathcal{N}(A) \setminus \{0\}$ such that $\|h_{\max(k)}\|_1 \geq \|h_{-\max(k)}\|_1$. Let $h = \sum_{i=1}^p c_i u_i$, where $\{c_i\}_{i=1}^p$ is a non-negative and non-increasing sequence; $\{u_i\}_{i=1}^p$ are indicator vectors (defined at the beginning of this section) with different supports in \mathbb{R}^p . Then we have $\sum_{i=1}^k c_i \geq \sum_{i=k+1}^p c_i$. Hence, $\|h_{-\max(a)}\|_\infty = c_{a+1} \leq \frac{\sum_{i=1}^a c_i}{a} =$

$\frac{\|h_{\max(a)}\|_1}{a}$ and

$$\begin{aligned} \|h_{-\max(a)}\|_1 &= \sum_{i=a+1}^k c_i + \sum_{i=k+1}^p c_i \leq \frac{k-a}{k} \sum_{i=1}^k c_i + \sum_{i=1}^k c_i \leq \frac{k-a}{a} \sum_{i=1}^a c_i + \frac{k}{a} \sum_{i=1}^a c_i \\ &= \frac{2k-a}{a} \|h_{\max(a)}\|_1. \end{aligned}$$

We set $\alpha = \frac{\|h_{\max(a)}\|_1}{a}$, $k_1 = a$, $k_2 = 2k - a$, It then follows from Lemma 1.1.1 that there exist $\{u_i\}_{i=1}^N, \{\lambda_i\}_{i=1}^N$ such that

$$h_{-\max(a)} = \sum_{i=1}^N \lambda_i u_i, \quad \|u_i\|_0 \leq 2k - a, \quad 0 \leq \lambda_i \leq 1, \quad \sum_{i=1}^N \lambda_i = 1,$$

$$\|h_{-\max(a)}\|_1 = \|u_i\|_1, \quad \|u\|_\infty \leq \frac{\|h_{\max(a)}\|_1}{a}.$$

Thus,

$$\begin{aligned} |\langle A(h_{\max(a)}), A(h_{-\max(a)}) \rangle| &\leq \sum_{i=1}^N \lambda_i |\langle A(h_{\max(a)}), A(u_i) \rangle| \\ &\leq \sum_{i=1}^N \lambda_i \theta_{a,2k-a}^A \|h_{\max(a)}\|_2 \|u_i\|_2 \\ &\leq \theta_{a,2k-a}^A \sqrt{2k-a} \|h_{\max(a)}\|_2 \cdot \frac{\|h_{\max(a)}\|_1}{a} \\ &\leq \theta_{a,2k-a}^A \sqrt{\frac{2k-a}{a}} \|h_{\max(a)}\|_2^2. \end{aligned} \tag{A.17}$$

On the other hand, Lemma A.1.7 yields

$$\theta_{a,2k-a} \leq \sqrt{\frac{2k-a}{\min\{b, 2k-a\}}} \theta_{a, \min\{b, 2k-a\}} \leq \max \left\{ \sqrt{\frac{2k-a}{b}}, 1 \right\} \theta_{a,b}.$$

Hence,

$$\begin{aligned}
0 &= |\langle A(h_{\max(a)}), A(h) \rangle| \geq |\langle A(h_{\max(a)}), A(h_{\max(a)}) \rangle| - |\langle A(h_{\max(a)}), A(h_{-\max(a)}) \rangle| \\
&\geq (1 - \delta_a^A) \|h_{\max(a)}\|_2^2 - \theta_{a,2k-a}^A \sqrt{\frac{2k-a}{a}} \|h_{\max(a)}\|_2^2 \\
&\geq (1 - \delta_a^A - \max \left\{ \frac{2k-a}{\sqrt{ab}}, \sqrt{\frac{2k-a}{a}} \right\} \theta_{a,b}^A) \|h_{\max(a)}\|_2^2 \\
&= (1 - \delta_a^A - C_{a,b,k} \theta_{a,b}^A) \|h_{\max(a)}\|_2^2
\end{aligned}$$

which contradicts the fact that $h \neq 0$ and $\delta_a^A + C_{a,b,k} \theta_{a,b}^A < 1$ and finished the proof of Theorem 1.2.2 and Part 2 of Theorem 1.1.1.

Part 3. If $\delta_{tk}^A < \sqrt{(t-1)/t}$.

First, we assume that tk is an integer. By the Null Space Property (Lemma A.1.2), we only need to check for all $h \in \mathcal{N}(A) \setminus \{0\}$, $\|h_{\max(k)}\|_1 < \|h_{-\max(k)}\|_1$. Suppose there exists $h \in \mathcal{N}(A) \setminus \{0\}$, such that $\|h_{\max(k)}\|_1 \geq \|h_{-\max(k)}\|_1$. Set $\alpha = \|h_{\max(k)}\|_1/k$. We divide $h_{-\max(k)}$ into two parts, $h_{-\max(k)} = h^{(1)} + h^{(2)}$, where

$$h^{(1)} = h_{-\max(k)} \cdot \mathbf{1}_{\{i \mid |h_{-\max(k)}(i)| > \alpha/(t-1)\}},$$

$$h^{(2)} = h_{-\max(k)} \cdot \mathbf{1}_{\{i \mid |h_{-\max(k)}(i)| \leq \alpha/(t-1)\}}.$$

Then $\|h^{(1)}\|_1 \leq \|h_{-\max(k)}\|_1 \leq \alpha k$. Denote $|\text{supp}(h^{(1)})| = \|h^{(1)}\|_0 = m$. Since all non-zero entries of $h^{(1)}$ have magnitude larger than $\alpha/(t-1)$, we have

$$\alpha k \geq \|h^{(1)}\|_1 = \sum_{i \in \text{supp}(h^{(1)})} |h^{(1)}(i)| \geq \sum_{i \in \text{supp}(h^{(1)})} \alpha/(t-1) = m\alpha/(t-1).$$

Namely $m \leq k(t-1)$. In addition we have

$$\begin{aligned} \|h^{(2)}\|_1 &= \|h_{-\max(k)}\|_1 - \|h^{(1)}\|_1 \leq k\alpha - \frac{m\alpha}{t-1} = (k(t-1) - m) \cdot \frac{\alpha}{t-1}, \\ \|h^{(2)}\|_\infty &\leq \frac{\alpha}{t-1}. \end{aligned} \quad (\text{A.18})$$

We now apply Lemma 1.1.1 with $s = k(t-1) - m$. Then $h^{(2)}$ can be expressed as a convex combination of sparse vectors: $h^{(2)} = \sum_{i=1}^N \lambda_i u_i$, where u_i is $(k(t-1) - m)$ -sparse and

$$\|u_i\|_1 = \|h^{(2)}\|_1, \quad \|u_i\|_\infty \leq \frac{\alpha}{(t-1)}, \quad \text{supp}(u_i) \subseteq \text{supp}(h^{(2)}). \quad (\text{A.19})$$

Hence,

$$\begin{aligned} \|u_i\|_2 &\leq \sqrt{\|u_i\|_0} \|u_i\|_\infty \leq \sqrt{k(t-1) - m} \|u_i\|_\infty \\ &\leq \sqrt{k(t-1)} \|u_i\|_\infty \leq \sqrt{k/(t-1)} \alpha. \end{aligned} \quad (\text{A.20})$$

Now we suppose $\mu \geq 0, c \geq 0$ are to be determined. Denote $\beta_i = h_{\max(k)} + h^{(1)} + \mu u_i$, then

$$\begin{aligned} \sum_{j=1}^N \lambda_j \beta_j - c\beta_i &= h_{\max(k)} + h^{(1)} + \mu h^{(2)} - c\beta_i \\ &= (1 - \mu - c)(h_{\max(k)} + h^{(1)}) - c\mu u_i + \mu h. \end{aligned} \quad (\text{A.21})$$

Since $h_{\max(k)}, h^{(1)}, u_i$ are k -, m -, $(k(t-1) - m)$ -sparse respectively, $\beta_i = h_{\max(k)} + h^{(1)} + \mu u_i, \sum_{j=1}^N \lambda_j \beta_j - c\beta_i - \mu h = (1 - \mu - c)(h_{\max(k)} + h^{(1)}) - c\mu u_i$ are all tk -sparse vectors.

We can check the following identity in ℓ_2 norm,

$$\begin{aligned} & \sum_{i=1}^N \lambda_i \|A(\sum_{j=1}^N \lambda_j \beta_j - c\beta_i)\|_2^2 + (1-2c) \sum_{1 \leq i < j \leq N} \lambda_i \lambda_j \|A(\beta_i - \beta_j)\|_2^2 \\ &= \sum_{i=1}^N \lambda_i (1-c)^2 \|A\beta_i\|_2^2. \end{aligned} \tag{A.22}$$

Since $Ah = 0$ and (A.21), we have $A(\sum_{j=1}^N \lambda_j \beta_j - c\beta_i) = A((1-\mu-c)(h_{\max(k)} + h^{(1)}) - c\mu u_i)$. Set $c = 1/2$, $\mu = \sqrt{t(t-1)} - (t-1)$, let the left hand side of (A.22) minus the right hand side, we get

$$\begin{aligned} 0 &\leq (1 + \delta_{tk}^A) \cdot \sum_{i=1}^N \lambda_i \left((1-\mu-c)^2 \|h_{\max(k)} + h^{(1)}\|_2^2 + c^2 \mu^2 \|u_i\|_2^2 \right) \\ &\quad - (1 - \delta_{tk}^A) \cdot \sum_{i=1}^N \lambda_i (1-c)^2 \left(\|h_{\max(k)} + h^{(1)}\|_2^2 + \mu^2 \|u_i\|_2^2 \right) \\ &= \sum_{i=1}^N \lambda_i \left[\left((1 + \delta_{tk}^A) \left(\frac{1}{2} - \mu \right)^2 - (1 - \delta_{tk}^A) \cdot \frac{1}{4} \right) \cdot \|h_{\max(k)} + h^{(1)}\|_2^2 + \frac{1}{2} \delta_{tk}^A \mu^2 \|u_i\|_2^2 \right] \\ &\leq \sum_{i=1}^N \lambda_i \|h_{\max(k)} + h^{(1)}\|_2^2 \cdot \left[(\mu^2 - \mu) + \delta_{tk}^A \left(\frac{1}{2} - \mu + \left(1 + \frac{1}{2(t-1)}\right) \mu^2 \right) \right] \\ &= \|h_{\max(k)} + h^{(1)}\|_2^2 \cdot \left[\delta_{tk}^A \left((2t-1)t - 2t\sqrt{t(t-1)} \right) \right. \\ &\quad \left. - \left((2t-1)\sqrt{t(t-1)} - 2t(t-1) \right) \right] < 0. \end{aligned}$$

We used the fact that

$$\begin{aligned} \delta_{tk}^A &< \sqrt{(t-1)/t}, \\ \|u_i\|_2 &\leq \sqrt{k/(t-1)} \alpha \leq \frac{\|h_{\max(k)}\|_2}{\sqrt{(t-1)}} \leq \frac{\|h_{\max(k)} + h^{(1)}\|_2}{\sqrt{t-1}} \end{aligned}$$

above. This is a contradiction.

When tk is not an integer, note $t' = \lceil tk \rceil / k$, then $t' > t$, $t'k$ is an integer,

$$\delta_{t'k} = \delta_{tk} < \sqrt{\frac{t-1}{t}} < \sqrt{\frac{t'k-1}{t'k}},$$

which can be deduced to the former case. Hence we finished the proof. \square

A.1.3 Proof of Theorems 1.2.1 and 1.2.3.

Suppose $h = \hat{\beta} - \beta$, where $\hat{\beta}$ may be $\hat{\beta}^{\ell_2}$ or $\hat{\beta}^{DS}$ depending on specific scenarios. For all the proofs, we will need a widely used fact. The readers may see Cai et al. (2009), Candès and Tao (2007), Candès et al. (2006), Donoho and Huo (2001) for details:

$$\|h_{-\max(k)}\|_1 \leq \|h_{\max(k)}\|_1 + 2\|\beta_{-\max(k)}\|_1. \quad (\text{A.23})$$

Again, the proof of Theorems 1.2.1 and 1.2.3 shall be divided into three parts: $\delta_k^A < 1/3$, $\delta_k^A + \theta_{k,k}^A < 1$ (or $\delta_a^A + C_{a,b,k}\theta_{a,b}^A < 1$) and $\delta_{tk}^A < \sqrt{(t-1)/t}$.

Part 1. If $\delta_k^A < 1/3$ for some $k \geq 2$.

We first prove the inequality for β^{ℓ_2} (1.10). Denote $h = \hat{\beta}^{\ell_2} - \beta$, then h can be written as $h = \sum_{i=1}^m a_i u_i$, where $\{u_i\}_{i=1}^p$ are indicator vectors with different support in \mathbb{R}^p ; $\{a_i\}_{i=1}^p$ is a non-negative and decreasing sequence. Then by (A.23) we have

$$\sum_{i=1}^k a_i + 2\|\beta_{-\max(k)}\|_1 \geq \sum_{i=k+1}^m a_i \quad (\text{A.24})$$

Apply Division Lemma A.1.1 by setting $a'_i = a_i + 2\|\beta_{-\max(k)}\|_1/k$, $i = 1, \dots, k$ and $a'_j = a_j$, $j > k + 1$, we can find $\{s_{ij}\}_{1 \leq i \leq k, 2k+1 \leq j \leq m}$ satisfying

$$\sum_{i=1}^k s_{ij} = a_j, \quad \forall 2k+1 \leq j \leq m, \quad (\text{A.25})$$

$$\frac{1}{k} \sum_{w=1}^k a_w + \frac{2\|\beta_{-\max(k)}\|_1}{k} \geq a_{k+i} + \sum_{j=2k+1}^m s_{ij}, \quad \forall 1 \leq i \leq k. \quad (\text{A.26})$$

We also know

$$\|Ah\| \leq \|A\beta - y\| + \|y - A\hat{\beta}\| \leq \epsilon + \eta. \quad (\text{A.27})$$

Similarly as Part 1 of Theorem 1.1.1, we finish the remaining part of proof for even or odd k separately.

1. When k is even, we define h_{11}, \dots, h_{32} as (A.7), similarly as (A.8) and by parallelogram equality, we get

$$\begin{aligned}
& \|A(-h_{11} + h_{22} + h_{32})\|^2 + \|A(-h_{12} + h_{21} + h_{31})\|^2 \\
&= \frac{1}{2} [\|A(-h_{11} - h_{12} + h_{21} + h_{22} + h_{31} + h_{32})\|^2 \\
&\quad + \|A(-h_{11} + h_{12} - h_{21} + h_{22} - h_{31} + h_{32})\|^2] \\
&= \frac{1}{2} \|A(2h_{11} + 2h_{12}) - Ah\|^2 + \frac{1}{4} \|A(-2h_{11} - 2h_{21} - 2h_{31})\|^2 \quad (\text{A.28}) \\
&\quad + \frac{1}{4} \|A(2h_{12} + 2h_{22} + 2h_{32})\|^2 - \frac{1}{8} \|A(2h)\|^2 \\
&= 2\|A(h_{11} + h_{12})\|^2 + \|A(h_{11} + h_{21} + h_{31})\|^2 \\
&\quad + \|A(h_{12} + h_{22} + h_{32})\|^2 - 2\langle Ah, A(h_{11} + h_{12}) \rangle
\end{aligned}$$

Let the right hand side of (A.28) minus the left hand side. Along with (A.9),

(A.10), one get

$$\begin{aligned}
0 &= RHS - LHS \\
&\geq 2(1 - \delta_k^A) \sum_{i=1}^k a_i^2 - 2\delta_k^A \sum_{i=1}^k a_i^2 - 2\delta_k^A \left(\sum_{i=k+1}^{2k} (a_i + \sum_{j=2k+1}^m s_{ij})^2 \right) \\
&\quad - 2\langle Ah, A(h_{11} + h_{12}) \rangle \\
&\geq 2(1 - 2\delta_k^A) \sum_{i=1}^k a_i^2 - 2\delta_k^A k \left(\frac{\sum_{i=1}^k a_i}{k} + \frac{2\|\beta_{-max(k)}\|_1}{k} \right)^2 \\
&\quad - 2(\epsilon + \eta) \sqrt{(1 + \delta_k^A) \sum_{i=1}^k a_i^2} \\
&\geq 2(1 - 2\delta_k^A) \sum_{i=1}^k a_i^2 - 2\delta_k^A \left(\sqrt{\sum_{i=1}^k a_i^2} + \frac{2\|\beta_{-max(k)}\|_1}{\sqrt{k}} \right)^2 \\
&\quad - 2(\epsilon + \eta) \sqrt{(1 + \delta_k^A) \sum_{i=1}^k a_i^2}
\end{aligned} \tag{A.29}$$

By (A.29) we can get an inequality of $\sqrt{\sum_{i=1}^k a_i^2}$:

$$\begin{aligned}
\sqrt{\sum_{i=1}^k a_i^2} &\leq \frac{\delta \frac{2\|\beta_{-max(k)}\|_1}{\sqrt{k}} + \frac{\epsilon + \eta}{2} \sqrt{1 + \delta}}{1 - 3\delta} \\
&\quad + \frac{\sqrt{(\delta \frac{2\|\beta_{-max(k)}\|_1}{\sqrt{k}} + \frac{\epsilon + \eta}{2} \sqrt{1 + \delta})^2 + (1 - 3\delta)\delta \|2\beta_{-max(k)}\|_1^2/k}}{1 - 3\delta} \\
&\leq \frac{\sqrt{1 + \delta}(\epsilon + \eta) + 2(2\delta + \sqrt{(1 - 3\delta)\delta})\|\beta_{-max(k)}\|_1/\sqrt{k}}{1 - 3\delta}
\end{aligned} \tag{A.30}$$

Finally, by Lemma A.1.5,

$$\sum_{i=k+1}^m a_i^2 \leq \left(\sqrt{\sum_{i=1}^k a_i^2} + \frac{2\|\beta_{-max(k)}\|_1}{\sqrt{k}} \right)^2$$

Then

$$\begin{aligned}
\|h\|_2 &= \sqrt{\sum_{i=1}^m a_i^2} \\
&\leq \sqrt{\sum_{i=1}^k a_i^2 + \left(\sqrt{\sum_{i=1}^k a_i^2} + \frac{2\|\beta_{-max(k)}\|_1}{\sqrt{k}}\right)^2} \leq \sqrt{2\sum_{i=1}^k a_i^2 + \frac{2\|\beta_{-max(k)}\|_1}{\sqrt{k}}} \\
&\leq \frac{\sqrt{2(1+\delta)}}{1-3\delta}(\epsilon + \eta) + \frac{2\sqrt{2}(2\delta + \sqrt{(1-3\delta)\delta}) + 2(1-3\delta)}{1-3\delta} \frac{\|\beta_{-max(k)}\|_1}{\sqrt{k}}
\end{aligned} \tag{A.31}$$

2. When k is odd, we use the definitions in (A.11). Similar equality as (A.28) holds as follows,

$$\begin{aligned}
&\|A(-h_{11} + h_{21} + h_{31})\|^2 + \|A(-h_{12} + h_{22} + h_{32})\|^2 + \|A(-h_{13} + h_{23} + h_{33})\|^2 \\
&= \|A(h_{12} + h_{13} + h_{21} + h_{31})\|^2 + \|A(h_{11} + h_{13} + h_{22} + h_{32})\|^2 \\
&\quad + \|A(h_{11} + h_{12} + h_{23} + h_{33})\|^2 + \|A(h_{11} + h_{12} + h_{13})\|^2 \\
&\quad - 2\langle A(h_{11} + h_{12} + h_{13}), Ah \rangle
\end{aligned}$$

By the method in the even case, we can still get the inequality (A.29). Hence we have the same estimation that finished the proof of (1.10). \square

The proof of the inequality for $\hat{\beta}^{DS}$ (1.13) is essentially the same. In this case, we shall use the inequalities

$$\|A^T Ah\|_\infty \leq \|A^T(A\beta - y)\|_\infty + \|A^T(y - A\hat{\beta})\|_\infty \leq \epsilon + \eta$$

and

$$|\langle Ah, Ah_{\max(k)} \rangle| = |h_{\max(k)}^T A^T Ah| \leq \|h_{\max(k)}\|_1 \|A^T Ah\|_\infty \leq \sqrt{k} \|h_{\max(k)}\|_2 (\epsilon + \eta)$$

in the calculation of (A.29).

Part 2. If $\delta_k^A + \theta_{k,k}^A < 1$ for some $k \geq 1$, or $\delta_a^A + C_{a,b,k} \theta_{a,b}^A < 1$ for some $1 \leq a \leq k, b \geq 1$.

Again, it suffices to prove Theorem 1.2.3 as Part 2 of Theorem 1.2.1 is a special case of Theorem 1.2.3.

We first prove the inequality for $\hat{\beta}^{\ell_2}$ (1.11). Set $h = \hat{\beta}^{\ell_2} - \beta$. By the boundedness of z and the definition of the feasible set for $\hat{\beta}$,

$$\|Ah\|_2 \leq \|Ah - y\|_2 + \|y - A\hat{\beta}\|_2 \leq \epsilon + \eta. \quad (\text{A.32})$$

On the other hand, suppose $h = \sum_{i=1}^p c_i u_i$, where $\{c_i\}_{i=1}^p$ are non-negative and non-decreasing, $\{u_i\}_{i=1}^p$ are indicator vectors with different supports. Then by (A.23) we have

$$\sum_{i=k+1}^m c_i \leq \sum_{i=1}^k c_i + 2\|\beta_{-\max(k)}\|_1. \quad (\text{A.33})$$

Hence, $\|h_{-\max(a)}\|_\infty = c_{a+1} \leq \frac{\sum_{i=1}^a c_i}{a} = \frac{\|h_{\max(a)}\|_1}{a} \leq \frac{\|h_{\max(a)}\|_1}{a} + \frac{2\|\beta_{-\max(k)}\|_1}{2k-a}$ and

$$\begin{aligned} \|h_{-\max(a)}\|_1 &= \sum_{i=a+1}^k c_i + \sum_{i=k+1}^p c_i \leq \frac{k-a}{k} \sum_{i=1}^k c_i + \sum_{i=1}^k c_i + 2\|\beta_{-\max(k)}\|_1 \\ &\leq \frac{k-a}{a} \sum_{i=1}^a c_i + \frac{k}{a} \sum_{i=1}^a c_i + 2\|\beta_{-\max(k)}\|_1 \\ &= \frac{2k-a}{a} \|h_{\max(a)}\|_1 + 2\|\beta_{-\max(k)}\|_1. \end{aligned}$$

Now set $\lambda = \frac{\|h_{\max(a)}\|_1}{a} + \frac{2\|\beta_{-\max(k)}\|_1}{2k-a}$, $k_1 = a$, and $k_2 = 2k - a$. Lemma 1.1.1 then yields that there exist $\{u_i\}_{i=1}^N, \{\lambda_i\}_{i=1}^N$ such that

$$h_{-\max(a)} = \sum_{i=1}^N \lambda_i u_i, \quad \|u_i\|_0 \leq 2k - a, \quad 0 \leq \lambda_i \leq 1, \quad \sum_{i=1}^N \lambda_i = 1,$$

$$\|h_{-\max(a)}\|_1 = \|u_i\|_1, \quad \|u_i\|_\infty \leq \frac{\|h_{\max}\|_1}{a} + \frac{2\|\beta_{-\max(k)}\|_1}{2k-a}.$$

Thus,

$$\begin{aligned}
|\langle A(h_{\max(a)}), A(h_{-\max(a)}) \rangle| &\leq \sum_{i=1}^N \lambda_i |\langle A(h_{\max(a)}), Au_i \rangle| \leq \sum_{i=1}^N \lambda_i \theta_{a,2k-a}^A \|h_{\max(a)}\|_2 \|u_i\|_2 \\
&\leq \theta_{a,2k-a}^A \sqrt{2k-a} \|h_{\max(a)}\|_2 \cdot \left(\frac{\|h_{\max(a)}\|_1}{a} + \frac{2\|\beta_{-\max(k)}\|_1}{2k-a} \right).
\end{aligned}$$

On the other hand,

$$|\langle Ah, Ah_{\max(a)} \rangle| \leq \|Ah\|_2 \|Ah_{\max(a)}\|_2 \leq (\epsilon + \eta) \sqrt{1 + \delta} \|h_{\max(a)}\|_2. \quad (\text{A.34})$$

Now we denote $\theta_{a,2k-a}$ as $\tilde{\theta}$, then

$$\begin{aligned}
&(\epsilon + \eta) \sqrt{1 + \delta} \|h_{\max(a)}\|_2 \\
&\geq |\langle Ah, Ah_{\max(a)} \rangle| \geq \|Ah_{\max(a)}\|_2^2 - |\langle Ah_{-\max(a)}, Ah_{\max(a)} \rangle| \\
&\geq (1 - \delta) \|h_{\max(a)}\|_2^2 - \tilde{\theta} \|h_{\max(a)}\|_2 \cdot \sqrt{2k-a} \left(\frac{\|h_{\max(a)}\|_1}{a} + \frac{2\|\beta_{-\max(k)}\|_1}{2k-a} \right) \\
&\geq (1 - \delta - \sqrt{\frac{2k-a}{a}} \tilde{\theta}) \|h_{\max(a)}\|_2^2 - \tilde{\theta} \|h_{\max(a)}\|_2 \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{2k-a}}.
\end{aligned}$$

Hence

$$\|h_{\max(a)}\|_2 \leq \frac{\sqrt{1 + \delta}(\epsilon + \eta)}{1 - \delta - \sqrt{(2k-a)/a}\tilde{\theta}} + \frac{\tilde{\theta}}{1 - \delta - \sqrt{(2k-a)/a}\tilde{\theta}} \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{2k-a}}. \quad (\text{A.35})$$

Applying Lemma A.1.5 with $\alpha = 2$ and $\lambda = 2\|h_{-\max(k)}\|_1$ yields

$$\begin{aligned}
\|h\|_2 &= \sqrt{\sum_{i=1}^k c_i^2 + \sum_{i=k+1}^p c_i^2} \leq \sqrt{\sum_{i=1}^k c_i^2 + \left(\sqrt{\sum_{i=1}^k c_i^2} + \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k}}\right)^2} \\
&\leq \sqrt{2\sum_{i=1}^k c_i^2 + \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k}}} \leq \sqrt{\frac{2k}{a}\sum_{i=1}^a c_i^2 + \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k}}} \\
&\leq \frac{\sqrt{2(1+\delta)k/a(\epsilon+\eta)}}{1-\delta-\sqrt{(2k-a)/a\tilde{\theta}}} \\
&\quad + \left(\frac{\sqrt{2k/a\tilde{\theta}}}{1-\delta-\sqrt{(2k-a)/a\tilde{\theta}}}\frac{2}{\sqrt{2k-a}} + \frac{2}{\sqrt{k}}\right)\|\beta_{-\max(k)}\|_1.
\end{aligned}$$

Finally, it follows from Lemma A.1.7 that

$$\begin{aligned}
\tilde{\theta} = \theta_{a,2k-a} &\leq \sqrt{\frac{2k-a}{\min\{b, 2k-a\}}}\theta_{a,\min\{b, 2k-a\}} \leq \max\left\{\sqrt{\frac{2k-a}{b}}, 1\right\}\theta_{a,b} \\
&= \sqrt{\frac{a}{2k-a}}C_{a,b,k}\theta_{a,b}.
\end{aligned}$$

So $\|h\|_2 \leq \frac{\sqrt{2(1+\delta)k/a(\epsilon+\eta)}}{1-\delta-C_{a,b,k}\theta} + 2\|\beta_{-\max(k)}\|_1 \left(\frac{\sqrt{2k}C_{a,b,k}\theta}{(1-\delta-C_{a,b,k}\theta)(2k-a)} + \frac{1}{\sqrt{k}}\right)$, which finishes the proof of 1.2.1.

The proof for $\hat{\beta}^{DS}$ is basically the same, where we only need to use the inequalities $\|A^T Ah\|_\infty \leq \|A^T(A\beta - y)\|_\infty + \|A^T(y - A\hat{\beta})\|_\infty \leq (\epsilon + \eta)$ and

$$|\langle Ah, Ah_{\max(a)} \rangle| = |h_{\max(a)}^T A^T Ah| \leq \|h_{\max(a)}\|_1 \|A^T Ah\|_\infty \leq \sqrt{a}\|h_{\max(a)}\|_2(\epsilon + \eta)$$

instead of (A.32) and (A.34).

Part 3. If $\delta_{tk}^A < \sqrt{(t-1)/t}$ for some $t \geq 4/3$.

We first prove the inequality on $\hat{\beta}^{\ell_2}$ (1.12). Similarly to the proof of Theorem

1.1.1, we assume that tk is an integer at first. Besides,

$$\|Ah\|_2 \leq \|y - A\beta\|_2 + \|A\hat{\beta}^{\ell_2} - y\|_2 \leq \epsilon + \eta. \quad (\text{A.36})$$

Define $\alpha = (\|h_{\max(k)}\|_1 + 2\|\beta_{-\max(k)}\|_1)/k$. Similarly as the proof of Theorem 1.1.1, we divide $h_{-\max(k)}$ into two parts, $h_{-\max(k)} = h^{(1)} + h^{(2)}$, where

$$h^{(1)} = h_{-\max(k)} \cdot \mathbf{1}_{\{|i||h_{-\max(k)}(i)| > \alpha/(t-1)\}}, \quad h^{(2)} = h_{-\max(k)} \cdot \mathbf{1}_{\{|i||h_{-\max(k)}(i)| \leq \alpha/(t-1)\}}.$$

Then $\|h^{(1)}\|_1 \leq \|h_{-\max(k)}\|_1 \stackrel{(\text{A.23})}{\leq} \alpha k$. Denote $|\text{supp}(h^{(1)})| = \|h^{(1)}\|_0 = m$. Since all non-zero entries of $h^{(1)}$ have magnitude larger than $\alpha/(t-1)$, we have

$$\alpha k \geq \|h^{(1)}\|_1 = \sum_{i \in \text{supp}(h^{(1)})} |h^{(1)}(i)| \geq \sum_{i \in \text{supp}(h^{(1)})} \alpha/(t-1) = m\alpha/(t-1).$$

Namely $m \leq k(t-1)$. Hence, (A.18) still holds. Besides, $\|h_{\max(k)} + h^{(1)}\|_0 = k + m \leq tk$, we have

$$\begin{aligned} \langle A(h_{\max(k)} + h^{(1)}), Ah \rangle &\leq \|A(h_{\max(k)} + h^{(1)})\|_2 \|Ah\|_2 \\ &\leq \sqrt{1 + \delta} \|h_{\max(k)} + h^{(1)}\|_2 (\epsilon + \eta). \end{aligned} \quad (\text{A.37})$$

Again by (A.18), we apply Lemma 1.1.1 by setting $s = k(t-1) - m$, we can express $h^{(2)}$ as a weighted mean: $h^{(2)} = \sum_{i=1}^N \lambda_i u_i$, where u_i is $(k(t-1) - m)$ -sparse and (A.19) still holds. Hence,

$$\|u_i\|_2 \leq \sqrt{\|u_i\|_0} \|u_i\|_\infty \leq \sqrt{k(t-1) - m} \|u_i\|_\infty \leq \sqrt{k(t-1)} \|u_i\|_\infty \leq \sqrt{k/(t-1)} \alpha.$$

Now we suppose $1 \geq \mu \geq 0, c \geq 0$ are to be determined. Denote $\beta_i = h_{\max(k)} + h^{(1)} + \mu u_i$, then we still have (A.21). Similarly to the proof of Theorem 1.1.1, since

$h_{\max(k)}, h^{(1)}, u_i$ are k -, m -, $(k(t-1) - m)$ -sparse vectors, respectively, we know $\beta_i = h_{\max(k)} + h^{(1)} + \mu u_i$, $\sum_{j=1}^N \lambda_j \beta_j - c\beta_i - \mu h = (1 - \mu - c)(h_{\max(k)} + h^{(1)}) - c\mu u_i$ are all tk sparse vectors.

Suppose $x = \|h_{\max(k)} + h^{(1)}\|_2$, $P = \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k}}$, then

$$\begin{aligned} \|u_i\|_2 &\leq \sqrt{k/(t-1)}\alpha \leq \frac{\|h_{\max(k)}\|_2}{\sqrt{(t-1)}} + \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k(t-1)}} \\ &\leq \frac{\|h_{\max(k)} + h^{(1)}\|_2}{\sqrt{t-1}} + \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k(t-1)}} = \frac{x + P}{\sqrt{t-1}}. \end{aligned}$$

We still use the ℓ_2 identity (A.22). Set $c = 1/2$, $\mu = \sqrt{t(t-1)} - (t-1)$ and take the difference of the left- and right-hand sides of (A.22), we get

$$\begin{aligned} 0 &= \sum_{i=1}^N \lambda_i \left\| A \left((h_{\max(k)} + h^{(1)} + \mu h^{(2)}) - \frac{1}{2}(h_{\max(k)} + h^{(1)} + \mu u_i) \right) \right\|_2^2 - \sum_{i=1}^N \frac{\lambda_i}{4} \|A\beta_i\|_2^2 \\ &= \sum_{i=1}^N \lambda_i \left\| A \left(\left(\frac{1}{2} - \mu\right)(h_{\max(k)} + h^{(1)}) - \frac{\mu}{2}u_i + \mu h \right) \right\|_2^2 - \sum_{i=1}^N \frac{\lambda_i}{4} \|A\beta_i\|_2^2 \\ &= \sum_{i=1}^N \lambda_i \left\| A \left(\left(\frac{1}{2} - \mu\right)(h_{\max(k)} + h^{(1)}) - \frac{\mu}{2}u_i \right) \right\|_2^2 \\ &\quad + 2 \left\langle A \left(\left(\frac{1}{2} - \mu\right)(h_{\max(k)} + h^{(1)}) - \frac{\mu}{2}h^{(2)} \right), \mu Ah \right\rangle + \mu^2 \|Ah\|_2^2 - \sum_{i=1}^N \frac{\lambda_i}{4} \|A\beta_i\|_2^2 \\ &= \sum_{i=1}^N \lambda_i \left\| A \left(\left(\frac{1}{2} - \mu\right)(h_{\max(k)} + h^{(1)}) - \frac{\mu}{2}u_i \right) \right\|_2^2 \\ &\quad + \mu(1 - \mu) \langle A(h_{\max(k)} + h^{(1)}), Ah \rangle - \sum_{i=1}^N \frac{\lambda_i}{4} \|A\beta_i\|_2^2. \end{aligned}$$

Now since β_i , $(\frac{1}{2} - \mu)(h_{\max(k)} + h^{(1)}) - \frac{\mu}{2}u_i$ are all tk -sparse vectors, we apply the

definition of δ_{ik}^A and also (A.37) to get

$$\begin{aligned}
0 &\leq (1+\delta) \sum_{i=1}^N \lambda_i \left(\left(\frac{1}{2} - \mu\right)^2 \|h_{\max(k)} + h^{(1)}\|_2^2 + \frac{\mu^2}{4} \|u_i\|_2^2 \right) \\
&\quad + \mu(1-\mu)\sqrt{1+\delta} \|h_{\max(k)} + h^{(1)}\|_2 (\epsilon + \eta) \\
&\quad - (1-\delta) \sum_{i=1}^N \frac{\lambda_i}{4} (\|h_{\max(k)} + h^{(1)}\|_2^2 + \mu^2 \|u_i\|_2^2) \\
&= \sum_{i=1}^N \lambda_i \left\{ \left((1+\delta)\left(\frac{1}{2} - \mu\right)^2 - (1-\delta) \cdot \frac{1}{4} \right) \cdot \|h_{\max(k)} + h^{(1)}\|_2^2 + \frac{1}{2} \delta \mu^2 \|u_i\|_2^2 \right\} \\
&\quad + \mu(1-\mu)\sqrt{1+\delta} \|h_{\max(k)} + h^{(1)}\|_2 (\epsilon + \eta) \\
&\leq \left[(\mu^2 - \mu) + \delta \left(\frac{1}{2} - \mu + \left(1 + \frac{1}{2(t-1)}\right) \mu^2 \right) \right] x^2 \\
&\quad + \left[\mu(1-\mu)\sqrt{1+\delta}(\epsilon + \eta) + \frac{\delta \mu^2 P}{t-1} \right] x + \frac{\delta \mu^2 P^2}{2(t-1)} \\
&= -t \left((2t-1) - 2\sqrt{t(t-1)} \right) \left(\sqrt{\frac{t-1}{t}} - \delta \right) x^2 \\
&\quad + \left[\mu^2 \sqrt{\frac{t}{t-1}} \cdot \sqrt{1+\delta}(\epsilon + \eta) + \frac{\delta \mu^2 P}{t-1} \right] x + \frac{\delta \mu^2 P^2}{2(t-1)} \\
&= \frac{\mu^2}{t-1} \left[-t \left(\sqrt{\frac{t-1}{t}} - \delta \right) x^2 + \left(\sqrt{t(t-1)}(1+\delta)(\epsilon + \eta) + \delta P \right) x + \frac{\delta P^2}{2} \right],
\end{aligned} \tag{A.38}$$

which is an second-order inequality for x . By solving this inequality we get

$$\begin{aligned}
x &\leq \left\{ \left(\sqrt{t(t-1)}(1+\delta)(\epsilon + \eta) + \delta P \right) + \left[\left(\sqrt{t(t-1)}(1+\delta)(\epsilon + \eta) + \delta P \right)^2 \right. \right. \\
&\quad \left. \left. + 2t(\sqrt{(t-1)/t} - \delta)\delta P^2 \right]^{1/2} \right\} \cdot \left(2t(\sqrt{(t-1)/t} - \delta) \right)^{-1} \\
&\leq \frac{\sqrt{t(t-1)}(1+\delta)}{t(\sqrt{(t-1)/t} - \delta)} (\epsilon + \eta) + \frac{2\delta + \sqrt{2t(\sqrt{(t-1)/t} - \delta)\delta}}{2t(\sqrt{(t-1)/t} - \delta)} P.
\end{aligned}$$

Finally, note that $\|h_{-\max(k)}\|_1 \leq \|h_{\max(k)}\|_1 + P\sqrt{k}$, by Lemma A.1.5, we obtain

$\|h_{-\max(k)}\|_2 \leq \|h_{\max(k)}\|_2 + P$, so

$$\begin{aligned}
\|h\|_2 &= \sqrt{\|h_{\max(k)}\|_2^2 + \|h_{-\max(k)}\|_2^2} \leq \sqrt{\|h_{\max(k)}\|_2^2 + (\|h_{\max(k)}\|_2 + P)^2} \\
&\leq \sqrt{2\|h_{\max(k)}\|_2^2} + P \leq \sqrt{2}x + P \\
&\leq \frac{\sqrt{2t(t-1)(1+\delta)}}{t(\sqrt{(t-1)/t} - \delta)}(\epsilon + \eta) + \left(\frac{\sqrt{2}\delta + \sqrt{t(\sqrt{(t-1)/t} - \delta)\delta}}{t(\sqrt{(t-1)/t} - \delta)} + 1 \right) \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k}} \\
&= \frac{\sqrt{2(1+\delta)}}{1 - \sqrt{t/(t-1)}\delta}(\epsilon + \eta) + \left(\frac{\sqrt{2}\delta + \sqrt{t(\sqrt{(t-1)/t} - \delta)\delta}}{t(\sqrt{(t-1)/t} - \delta)} + 1 \right) \frac{2\|\beta_{-\max(k)}\|_1}{\sqrt{k}},
\end{aligned}$$

which finished the proof.

When tk is not an integer, again we define $t' = \lceil tk \rceil / k$, then $t' > t$ and $\delta_{t'k}^A = \delta_{tk}^A < \sqrt{\frac{t-1}{t}} < \sqrt{\frac{t'-1}{t'}}$. We can prove the result by working on $\delta_{t'k}^A$.

For the inequality on $\hat{\beta}^{DS}$ (1.15), the proof is similar. Define $h = \hat{\beta}^{DS} - \beta$. We have the following inequalities

$$\|A^T Ah\|_\infty \leq \|A^T(A\hat{\beta}^{\ell_2} - y)\|_\infty + \|A^T(y - A\beta)\|_\infty \leq \eta + \epsilon,$$

$$\begin{aligned}
\langle A(h_{\max(k)} + h^{(1)}), Ah \rangle &= \langle h_{\max(k)} + h^{(1)}, A^T Ah \rangle \\
&\leq \|h_{\max(k)} + h^{(1)}\|_1(\epsilon + \eta) \leq \sqrt{tk}(\epsilon + \eta)\|h_{\max(k)+h^{(1)}}\|_2,
\end{aligned} \tag{A.39}$$

instead of (A.36) and (A.37). We can prove (1.15) basically the same as the proof above except that we use (A.39) instead of (A.37) when we go from the third term to the fourth term in (A.38). \square

A.1.4 Proof of Proposition 1.2.1.

By a small extension of Lemma 5.1 in Cai et al. (2009), we have $\|z\|_2 \leq \sigma\sqrt{n + 2\sqrt{n \log n}}$ with probability at least $1 - 1/n$; $\|A^T z\|_\infty \leq \sigma\sqrt{2(1 + \delta_1^A) \log p} \leq 2\sigma\sqrt{\log p}$ with probability at least $1 - 1/\sqrt{\pi \log p}$. Then the Proposition is immediately implied by

Theorem 1.2.1. \square

A.1.5 Proof of Proposition 1.2.2.

The proof of Proposition (1.2.2) is similar to that of Theorem 2.7 in Candès and Plan (2011).

First, as in the proof of Proposition 1.2.1, we have $\|A^T z\|_\infty \leq \lambda/2$ with probability at least $1/\sqrt{\pi \log n}$. In the rest proof, we will prove (1.19), (1.20) and (1.21) under different RIP conditions in the event that $\|A^T z\|_\infty \leq \lambda/2$. Define

$$K(\xi, \beta) = \gamma \|\xi\|_0 + \|A\beta - A\xi\|_2^2, \quad \gamma = \frac{\lambda^2}{8} = 2\sigma^2 \log p.$$

Let $\bar{\beta} = \arg \min_\xi K(\xi, \beta)$. Since $K(\bar{\beta}, \beta) \leq K(\beta, \beta)$, we have $\gamma \|\bar{\beta}\|_0 \leq \gamma \|\beta\|_0$, which means $\bar{\beta}$ is k -sparse.

With a small edition on Lemma 3.5 in Candès and Plan (2011), we can prove

$$\|A^T A(\bar{\beta} - \beta)\|_\infty \leq \lambda/2 \tag{A.40}$$

In fact, if (A.40) does not hold. Suppose $|(A^T A(\bar{\beta} - \beta))_i| > \lambda/2$, i.e. the absolute value of the i -th entry is greater than $\lambda/2$. We construct

$$\bar{\beta}' = \bar{\beta} - \alpha e_i, \quad \alpha = \frac{(A^T A(\bar{\beta} - \beta))_i}{\|Ae_i\|_2^2}.$$

Then

$$\|A(\bar{\beta}' - \beta)\|_2^2 = \|A(\bar{\beta} - \beta)\|_2^2 - 2\alpha \langle Ae_i, A(\bar{\beta} - \beta) \rangle + \alpha^2 \|Ae_i\|_2^2 = \|A(\bar{\beta} - \beta)\|_2^2 - \alpha^2 \|Ae_i\|_2^2$$

which yields

$$\begin{aligned}
K(\bar{\beta}', \beta) &\leq \gamma(\|\bar{\beta}\|_0 + 1) + \|A(\bar{\beta} - \beta)\|_2^2 - \alpha^2 \|Ae_i\|_2^2 \\
&\leq K(\bar{\beta}, \beta) + \gamma - \alpha^2 \|Ae_i\|_2^2 \\
&\leq K(\bar{\beta}, \beta) + \frac{\lambda^2}{8} - (A^T A(\bar{\beta} - \beta))_i^2 / \|Ae_i\|_2^2 \\
&< K(\bar{\beta}, \beta) + \frac{\lambda^2}{8} - (\lambda/2)^2 / (1 + \delta_1^A) \leq K(\bar{\beta}, \beta),
\end{aligned}$$

This is a contradiction to the assumption that $\bar{\beta}$ is the minimizer of $K(\xi, \beta)$, namely (A.40) holds. So we have

$$\|A^T(y - A\bar{\beta})\|_\infty \leq \|A^T(y - A\beta)\|_\infty + \|A^T A(\beta - \bar{\beta})\|_\infty \leq \lambda. \quad (\text{A.41})$$

So $\bar{\beta}$ is feasible in (1.2). Since $\|\bar{\beta}\|_0 \leq k$, we can apply Theorem 1.2.1 by plugging β by $\bar{\beta}$ and get

$$\|\hat{\beta} - \bar{\beta}\|_2 \leq \begin{cases} \frac{\sqrt{2\|\bar{\beta}\|_0}}{1-3\delta_k^A} 2\lambda, & \delta_k^A < 1/3; \\ \frac{\sqrt{2\|\bar{\beta}\|_0}}{1-\delta_k^A - \theta_{k,k}^A} 2\lambda, & \delta_k^A + \theta_{k,k}^A < 1; \\ \frac{\sqrt{2t\|\bar{\beta}\|_0}}{1-\sqrt{t/(t-1)}\delta_{tk}^A} 2\lambda, & \delta_{tk}^A < \sqrt{(t-1)/t}. \end{cases} \quad (\text{A.42})$$

Next, we prove $\delta_{2k}^A < 1$ under any of the three RIP conditions.

1. When $\delta_k^A < 1/3$, by Lemma A.1.6, we have

$$\delta_{2k} \leq 3\delta_k^A < 1.$$

2. When $\delta_k^A + \theta_{k,k}^A < 1/3$, by Lemma 1.1 in Candès and Tao (2005), we have

$$\delta_{2k} \leq \delta_k^A + \theta_{k,k}^A < 1.$$

3. When $\delta_{tk}^A < \sqrt{(t-1)/t}$, by Lemma A.1.6, we can see when $1 < t < 2$,

$$\delta_{2k}^A \leq \left(2 \frac{2k}{\lceil tk \rceil} - 1\right) \delta_{\lceil tk \rceil}^A \leq (4/t - 1) \delta_{tk}^A \leq \sqrt{t/(t-1)} \delta_{tk}^A.$$

When $t \geq 2$, $\delta_{2k}^A \leq \delta_{tk}^A$, which means

$$\delta_{2k}^A \leq \sqrt{t/(t-1)} \delta_{tk}^A < 1, \quad (\text{A.43})$$

whenever $t \geq 4/3$.

Finally, we finish the proof for the case where $\delta_{tk}^A < \sqrt{(t-1)/t}$. The other two cases can follow with small editions.

$$\|\bar{\beta} - \beta\|_2^2 \leq \frac{1}{1 - \delta_{2k}^A} \|A\bar{\beta} - A\beta\|_2^2 \leq \frac{1}{1 - \sqrt{t/(t-1)} \delta_{tk}^A} \|A\bar{\beta} - A\beta\|_2^2.$$

Hence,

$$\begin{aligned} \|\hat{\beta} - \beta\|_2^2 &\leq 2\|\hat{\beta} - \bar{\beta}\|_2^2 + 2\|\bar{\beta} - \beta\|_2^2 \\ &\leq \frac{16t\|\bar{\beta}\|_0\lambda^2}{(1 - \sqrt{t/(t-1)} \delta_{tk}^A)^2} + \frac{2}{1 - \sqrt{t/(t-1)} \delta_{tk}^A} \|A\bar{\beta} - A\beta\|_2^2 \\ &\leq \frac{128t}{(1 - \sqrt{t/(t-1)} \delta_{tk}^A)^2} K(\bar{\beta}, \beta). \end{aligned}$$

Suppose $\beta' = \sum_{i=1}^p \beta \cdot 1_{\{|\beta_i| > \mu\}}$, where $\mu = \sqrt{\frac{\gamma}{1 + \delta_k^A}}$. Then

$$\begin{aligned} K(\bar{\beta}, \beta) &\leq K(\beta', \beta) \leq \gamma \sum_{i=1}^p 1_{\{|\beta_i| > \mu\}} + \|A\beta' - A\beta\|_2^2 \\ &\leq \gamma \sum_{i=1}^p 1_{\{|\beta_i| > \mu\}} + (1 + \delta_k^A) \sum_{i=1}^p 1_{\{|\beta_i| \leq \mu\}} |\beta_i|^2 \\ &\leq \sum_{i=1}^p \min(\gamma, (1 + \delta_k^A) |\beta_i|^2) \leq 2 \log p \sum_{i=1}^p \min(\sigma^2, |\beta_i|^2). \end{aligned}$$

Therefore, we have proved (1.21) in the event that $\|A^T z\|_\infty \leq \lambda/2$. \square

A.1.6 Proof of Theorem 1.2.4.

Again, we divide the proof of Theorem 1.2.4 into three parts.

Part 1. “ $\delta_k^A < 1/3$ ” is sharp.

Note

$$\beta_1 = \text{diag}\left(\overbrace{\frac{1}{\sqrt{2k}}, \dots, \frac{1}{\sqrt{2k}}}^{2k}, 0, \dots, 0\right) \in \mathbb{R}^p$$

Suppose $H = (\mathbb{R}^p, \|\beta\|_2)$ is the Hilbert with inner product $\langle \cdot, \cdot \rangle$. Since $\|\beta_1\|_2 = 1$, we can extend β_1 into a basis $\{\beta_1, \dots, \beta_p\}$. Define $A : \mathbb{R}^p \rightarrow \mathbb{R}^p$ as

$$A\gamma = \sqrt{\frac{4}{3}} \sum_{i=2}^p a_i \beta_i \tag{A.44}$$

for all $\gamma = \sum_{i=1}^p a_i \beta_i$.

Then by Cauchy-Schwarz inequality, for all k -sparse vector γ , we have

$$|\langle \gamma, \beta_1 \rangle| = |\langle \gamma, \beta_1 \cdot \mathbf{1}_{\text{supp}(\gamma)} \rangle| \leq \|\gamma_{\max(k)}\|_2 \|\beta_1\|_2 \leq \sqrt{k \cdot \frac{1}{2k}} \|\beta_1\|_2 = \sqrt{\frac{1}{2}} \|\beta_1\|_2$$

$$\|A\gamma\|_2^2 = \frac{4}{3} \sum_{i=2}^p a_i^2 = \frac{4}{3} (\|\gamma\|_2^2 - a_1^2) = \frac{4}{3} (\|\gamma\|_2^2 - |\langle \gamma, \beta_1 \rangle|^2)$$

Thus,

$$\frac{2}{3} \|\gamma\|_2^2 \leq \|A\gamma\|_2^2 \leq \frac{4}{3} \|\gamma\|_2^2, \quad \delta_k^A \leq 1/3$$

Notice that

$$\|A(\overbrace{1, \dots, 1}^k, 0, \dots, 0)\|_2^2 = \frac{2}{3}k = \frac{2}{3} \|(\overbrace{1, \dots, 1}^k, 0, \dots, 0)\|_F^2$$

$$\|A(1, -1, 0, \dots, 0)\|_2^2 = \frac{8}{3} = \frac{4}{3} \|(1, -1, 0, \dots, 0)\|_F^2$$

we can conclude that $\delta_k^A = 1/3$. Finally, suppose

$$u = (\overbrace{1, 1, \dots, 1}^k, 0, \dots, 0) \in \mathbb{R}^p, \quad v = (\overbrace{0, \dots, 0}^k, \overbrace{-1, -1, \dots, -1}^k, 0, \dots, 0) \in \mathbb{R}^p$$

Then u, v are both matrices of rank k such that $Au = Av$. Therefore, it is impossible to recover both u and v only given (y, A) in both the noiseless and noisy case, which finishes the proof of Part 1 of Theorem 1.2.4.

Part 2. “ $\delta_{tk}^A + \theta_{k,k}^A < 1$ ” and “ $\delta_a^A + C_{a,b,k}\theta_{a,b}^A < 1$ ” are sharp.

Again, it suffices to prove Theorem 1.2.5 as Part 2 of Theorem 1.2.4 is a special case of Theorem 1.2.5.

Still, we define A as (A.44). The Cauchy-Schwarz Inequality yields that $|\langle \gamma, \beta_1 \rangle| \leq \|\beta_1 \cdot \mathbf{1}_{\text{supp}(\gamma)}\|_2 \|\gamma\|_2 \leq \sqrt{\frac{a}{2k}} \|\gamma\|_2$ for all a -sparse vector γ . Note that $\|A\gamma\|_2^2 = \sum_{i=2}^p c_i^2 = \frac{2}{2-a/(2k)} (\|\gamma\|_2^2 - c_1^2) = \frac{2}{2-a/(2k)} (\|\gamma\|_2^2 - |\langle \gamma, \beta_1 \rangle|^2)$. So

$$\left(1 - \frac{a/(2k)}{2 - a/(2k)}\right) \|\gamma\|_2^2 \leq \|A\gamma\|_2^2 \leq \left(1 + \frac{a/(2k)}{2 - a/(2k)}\right) \|\gamma\|_2^2 \quad \text{and} \quad \delta_a^A \leq \frac{a/(2k)}{2 - a/(2k)}.$$

Now we estimate $\theta_{a,b}^A$. For any a -sparse vector γ_1 and b -sparse vector $\gamma_2 \in \mathbb{R}^p$ with disjoint supports, write $\gamma_1 = \sum_{i=1}^p c_i \beta_i$ and $\gamma_2 = \sum_{i=1}^p d_i \beta_i$, we have $\frac{a/(2k)}{2-a/(2k)} \sum_{i=1}^p c_i d_i = \langle \gamma_1, \gamma_2 \rangle = 0$.

1. When $b \leq 2k - a$, The Cauchy-Schwarz Inequality yields that $|c_1| = |\langle \beta_1, \gamma_1 \rangle| \leq \sqrt{\frac{a}{2k}} \|\gamma_1\|_2$ and $|d_1| = |\langle \beta_1, \gamma_2 \rangle| \leq \sqrt{\frac{b}{2k}} \|\gamma_2\|_2$. So

$$\frac{2 - a/(2k)}{2} |\langle A\gamma_1, A\gamma_2 \rangle| = \left| \sum_{i=2}^p c_i d_i \right| = |-c_1 d_1| \leq \frac{\sqrt{ab}}{2k} \|\gamma_1\|_2 \|\gamma_2\|_2$$

and consequently $\theta_{a,b} \leq \frac{2}{2-a/(2k)} \cdot \frac{\sqrt{ab}}{2k}$. Hence

$$\delta_a^A + C_{a,b,k} \theta_{a,b}^A \leq \frac{a/(2k)}{2-a/(2k)} + \max \left\{ \frac{2k-a}{\sqrt{ab}}, \sqrt{\frac{2k-a}{a}} \right\} \cdot \frac{2}{2-a/(2k)} \frac{\sqrt{ab}}{2k} \leq 1.$$

2. When $b > 2k - a$, if $\gamma_1 = 0$ or $\gamma_2 = 0$, it is clear that $\langle A\gamma_1, A\gamma_2 \rangle = 0 \leq C \|\gamma_1\|_2 \|\gamma_2\|_2$ for any $C \geq 0$. Without loss of generality, we assume that γ_1 and γ_2 are non-zero and are normalized so that $\|\gamma_1\|_2 = \|\gamma_2\|_2 = 1$. Since γ_1 and γ_2 are a , b -sparse respectively and γ_1 and γ_2 have disjoint supports, it follows from the Cauchy-Schwarz Inequality that for all $\lambda \geq 0$, $|c_1| = |\langle \beta_1, \gamma_1 \rangle| \leq \sqrt{\frac{a}{2k}} \|\gamma_1\|_2 = \sqrt{\frac{a}{2k}}$ and

$$\begin{aligned} \left| d_1 \pm \sqrt{\frac{a}{2k-a}} c_1 \right| &= \left| \langle \beta_1, \gamma_2 \pm \sqrt{\frac{a}{2k-a}} \gamma_1 \rangle \right| \leq \left\| \gamma_2 \pm \sqrt{\frac{a}{2k-a}} \gamma_1 \right\|_2 \\ &= \sqrt{\frac{2k}{2k-a}}. \end{aligned}$$

Hence,

$$\begin{aligned} &\frac{2-a/(2k)}{2} |\langle A\gamma_1, A\gamma_2 \rangle| = \left| \sum_{i=2}^{mn} c_i d_i \right| = |-c_1 d_1| \\ &= |c_1| \cdot \left(\max\{|d_1 + \sqrt{\frac{a}{2k-a}} c_1|, |d_1 - \sqrt{\frac{a}{2k-a}} c_1|\} - \left| \sqrt{\frac{a}{2k-a}} c_1 \right| \right) \\ &\leq |c_1| \cdot \left(\sqrt{\frac{2k}{2k-a}} - \sqrt{\frac{a}{2k-a}} |c_1| \right) \\ &= -\sqrt{\frac{a}{2k-a}} \left(\sqrt{\frac{k}{2a}} - |c_1| \right)^2 + \frac{k}{2\sqrt{a(2k-a)}} \\ &\leq \frac{\sqrt{a(2k-a)}}{2k} \end{aligned}$$

where the last inequality follows from the facts that $|c_1| \leq \sqrt{a/(2k)}$ and $a \leq k$.

So $\theta_{a,b}^A \leq \frac{2}{2-a/(2k)} \cdot \frac{\sqrt{a(2k-a)}}{2k}$ and

$$\begin{aligned} & \delta_a^A + C_{a,b,k}\theta_{a,b}^A \\ & \leq \frac{a/(2k)}{2-a/(2k)} + \max \left\{ \frac{2k-a}{\sqrt{ab}}, \sqrt{\frac{2k-a}{a}} \right\} \cdot \frac{2}{2-a/(2k)} \frac{\sqrt{a(2k-a)}}{2k} \leq 1. \end{aligned}$$

To sum up, we have shown $\delta_a^A + C_{a,b,k}\theta_{a,b}^A \leq 1$. Furthermore, let

$$u = (\overbrace{1, \dots, 1}^k, 0, \dots) \quad \text{and} \quad v = (\overbrace{0, \dots, 0}^k, \overbrace{-1, \dots, -1}^k, 0, \dots),$$

so u and v are both k -sparse and $Au = Av$, since $A(u-v) = 0$. Suppose $y = Au = Av$, then the k -sparse signals u and v are not distinguishable based on (y, A) .

Thus, the general recovery cannot be done in both noiseless and noisy case. Finally, $\delta_a^A + C_{a,b,k}\theta_{a,b}^A < 1$ is impossible by Theorem 1.2.2, so we must have $\delta_a^A + C_{a,b,k}\theta_{a,b}^A = 1$.

Part 3. “ $\delta_{tk}^A < \sqrt{(t-1)/t}$ ” is sharp.

For any $\epsilon > 0$ and $k \geq 5/\epsilon$, suppose $p \geq 2tk$, $m' = ((t-1) + \sqrt{t(t-1)})k$, m is the largest integer strictly smaller than m' . Then $m < m'$ and $m' - m \leq 1$. Since $t \geq 4/3$, we have $m' \geq k$. Define

$$\beta_1 = \sqrt{k + \frac{mk^2}{m'^2}}^{-1} (\overbrace{1, \dots, 1}^k, \overbrace{-\frac{k}{m'}, \dots, -\frac{k}{m'}}^m, 0, \dots, 0) \in \mathbb{R}^p,$$

then $\|\beta_1\|_2 = 1$. We define linear map $A : \mathbb{R}^p \rightarrow \mathbb{R}^p$, such that for all $\beta \in \mathbb{R}^p$,

$$A\beta = \sqrt{1 + \sqrt{\frac{t-1}{t}}} (\beta - \langle \beta_1, \beta \rangle \beta_1).$$

Now for all $[tk]$ -sparse vector β ,

$$\begin{aligned}\|A\beta\|_2^2 &= \left(1 + \sqrt{\frac{t-1}{t}}\right) (\beta - \langle\beta_1, \beta\rangle\beta_1)^T (\beta - \langle\beta_1, \beta\rangle\beta_1) \\ &= \left(1 + \sqrt{\frac{t-1}{t}}\right) (\|\beta\|_2^2 - |\langle\beta_1, \beta\rangle|^2).\end{aligned}$$

Since β is $[tk]$ -sparse, by Cauchy-Schwarz Inequality,

$$\begin{aligned}0 &\leq |\langle\beta_1, \beta\rangle|^2 \leq \|\beta\|_2^2 \cdot \|\beta_1 \cdot \mathbf{1}_{\text{supp}(\beta)}\|_2^2 \leq \|\beta\|_2^2 \|\beta_{1, \max([tk])}\|_2^2 \\ &= \|\beta\|_2^2 \cdot \frac{m'^2 + k([tk] - k)}{m'^2 + mk} \leq \frac{m'^2 + k^2(t-1) + k}{m'^2 + m'k} \cdot \frac{1}{1 - \frac{k(m'-m)}{m'^2 + m'k}} \|\beta\|_2^2 \\ &= \frac{m'^2 + k^2(t-1)}{m'^2 + m'k} \cdot \frac{m'^2 + k^2(t-1) + k}{m'^2 + k^2(t-1)} \cdot \frac{1}{1 - \frac{k(m'-m)}{m'^2 + m'k}} \|\beta\|_2^2 \\ &= 2\sqrt{t-1}(\sqrt{t} - \sqrt{t-1}) \cdot \left(1 + \frac{1}{tk}\right) \cdot \frac{1}{1 - \frac{1}{2k}} \|\beta\|_2^2 \\ &\leq \left(2\sqrt{t(t-1)} - 2(t-1)\right) \cdot \left(1 + \frac{5}{2k}\right) \|\beta\|_2^2 \leq \left(2\sqrt{t(t-1)} - 2(t-1) + \frac{5}{2k}\right) \|\beta\|_2^2.\end{aligned}$$

We used the fact that $m' \geq k$, $0 < m' - m \leq 1$ and

$$\begin{aligned}\frac{m'^2 + k^2(t-1)}{m'^2 + m'k} &= \frac{\left((t-1) + \sqrt{t(t-1)}\right)^2 + t-1}{\left((t-1) + \sqrt{t(t-1)}\right)^2 + \left((t-1) + \sqrt{t(t-1)}\right)} \\ &= \frac{(t-1)\left(t-1+t+2\sqrt{t(t-1)}+1\right)}{\sqrt{t(t-1)}\left(\sqrt{t} + \sqrt{t-1}\right)^2} = \frac{2\sqrt{t-1}}{\sqrt{t} + \sqrt{t-1}} = 2\sqrt{t-1}\left(\sqrt{t} - \sqrt{t-1}\right)\end{aligned}$$

above. Hence,

$$\begin{aligned}\left(1 + \sqrt{\frac{t-1}{t}}\right) \|\beta\|_2^2 &\geq \|A\beta\|_2^2 \geq \left(1 - \sqrt{\frac{t-1}{t}} - \left(1 + \sqrt{\frac{t-1}{t}}\right) \frac{5}{2k}\right) \|\beta\|_2^2 \\ &\geq \left(1 - \sqrt{\frac{t-1}{t}} - \epsilon\right) \|\beta\|_2^2,\end{aligned}$$

which implies $\delta_{tk}^A \leq \sqrt{(t-1)/t} + \epsilon$.

Now we consider

$$\beta_0 = (\overbrace{1, \dots, 1}^k, 0, \dots, 0) \in \mathbb{R}^p,$$

$$\gamma_0 = (\overbrace{0, \dots, 0}^k, \overbrace{\frac{k}{m'}, \dots, \frac{k}{m'}}^m, 0, \dots, 0).$$

Note that $A\beta_1 = 0$, so $A\beta_0 = A\gamma_0$. Besides, β_0 is k -sparse and $\|\gamma_0\|_1 < \|\beta_0\|_1$.

- In the noiseless case, i.e. $y = A\beta_0$, the ℓ_1 minimization method (1.2) fails to exactly recover β_0 through y since $y = A\gamma_0$, but $\|\gamma_0\|_1 < \|\beta_0\|_1$.
- In the noisy case, i.e. $y = A\beta_0 + z$, assume that ℓ_1 minimization method (1.2) can stably recover β_0 with constraint \mathcal{B}_z . Suppose $\hat{\beta}_z$ is the solution of ℓ_1 minimization, then $\lim_{z \rightarrow 0} \hat{\beta}_z = \beta_0$. Note that $y - A(\hat{\beta}_z - \beta_0 + \gamma_0) = y - A\hat{\beta}_z \in \mathcal{B}_z$, by the definition of $\hat{\beta}_z$, we have $\|\hat{\beta}_z - \beta_0 + \gamma_0\|_1 \geq \|\hat{\beta}_z\|_1$. Let $z \rightarrow 0$, it contradicts that $\|\gamma_0\|_1 < \|\beta_0\|_1$. Therefore, ℓ_1 minimization method (1.2) fails to stably recover β_0 . \square

A.1.7 Proof of Proposition 1.4.1.

Based on Theorem 1.2.2,

$$\delta_{tk}^A + \frac{2k - tk}{tk} \theta_{tk,tk}^A < 1 \quad (\text{A.45})$$

is a sufficient condition for exact recovery of all k -sparse vectors. By Lemma A.1.8,

$\theta_{tk,tk}^A \leq 2\delta_{tk}^A$ when tk is even; $\theta_{tk,tk}^A \leq \frac{2tk}{\sqrt{(tk)^2 - 1}} \delta_{tk}^A$ when tk is odd. Hence,

$$\delta_{tk}^A + \frac{2k - tk}{tk} \theta_{tk,tk}^A \leq \frac{4 - t}{t} \delta_{tk}^A, \quad tk \text{ is even;}$$

$$\delta_{tk}^A + \frac{2k - tk}{tk} \theta_{tk,tk}^A \leq \left(1 + \frac{4k - 2tk}{\sqrt{(tk)^2 - 1}} \right) \delta_{tk}^A, \quad tk \text{ is odd.}$$

The proposition is implied by the inequalities above and (A.45). \square

A.1.8 Proof of Proposition 1.4.2.

The idea of the proof is quite similar to Theorems 1.2.4 and 1.2.5. Define

$$\gamma = \frac{1}{\sqrt{2k}}(\overbrace{1, \dots, 1}^{2k}, 0, \dots, 0),$$

$$A : \mathbb{R}^p \rightarrow \mathbb{R}^p$$

$$\beta \mapsto \frac{2}{\sqrt{4-t}}(\beta - \langle \beta, \gamma \rangle \gamma).$$

Now for all non-zero $\lceil tk \rceil$ -sparse vector $\beta \in \mathbb{R}^p$,

$$\|A\beta\|_2^2 = \frac{4}{4-t} \langle \beta - \langle \beta, \gamma \rangle \gamma, \beta - \langle \beta, \gamma \rangle \gamma \rangle = \frac{4}{4-t} (\|\beta\|_2^2 - \langle \beta, \gamma \rangle^2).$$

We can immediately see $\|A\beta\|_2^2 \leq (1 + t/(4-t))\|\beta\|_2^2$. On the other hand by Cauchy-Schwarz's inequality,

$$\langle \beta, \gamma \rangle^2 = \langle \beta, \gamma \cdot \mathbf{1}_{\text{supp}(\beta)} \rangle^2 \leq \|\beta\|_2^2 \left(\sum_{i \in \text{supp}(\beta)} \gamma_i^2 \right) \leq \|\beta\|_2^2 \cdot \frac{\lceil tk \rceil}{2k}.$$

For $k > 1/\epsilon$, we have

$$\|A\beta\|_2^2 \geq \frac{4}{4-t} \left(1 - \frac{\lceil tk \rceil}{2k}\right) \|\beta\|_2^2 \geq \frac{4}{4-t} \left(1 - \frac{tk}{2k} - \epsilon/2\right) \|\beta\|_2^2 > \left(1 - \frac{t}{4-t} - \epsilon\right) \|\beta\|_2^2.$$

Therefore, we must have $\delta_{tk}^A = \delta_{\lceil tk \rceil}^A < t/(4-t) + \epsilon$.

Finally, we define

$$\beta_0 = (\overbrace{1, \dots, 1}^k, 0, \dots, 0), \quad \beta'_0 = (\overbrace{0, \dots, 0}^k, \overbrace{-1, \dots, -1}^k, 0, \dots, 0).$$

Then β_0, β'_0 are both k -sparse, and $y = A\beta_0 = A\beta'_0$. There's no way to recover both β_0, β'_0 only from (y, A) . \square

A.1.9 Technical Lemmas

In this section, we collect all technical tools for the proof of main theorems in Chapter 1.

It is well known that for matrices X, B with the same size, $|\langle X, B \rangle| \leq \|X\|_F \|B\|_F$. The following lemma provides a stronger result given further constraint on matrix rank.

Lemma A.1.3. *Let $X \in \mathbb{R}^{m \times n}$ ($m \leq n$) be a matrix with singular values $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$, then for all $B \in \mathbb{R}^{m \times n}$ such that $\text{rank}(B) \leq r$, we have*

$$|\langle B, X \rangle| \leq \|B\|_F \sqrt{\sum_{i=1}^r \lambda_i^2}.$$

Proof of Lemma A.1.3 Since the rank of B is at most r , we can suppose B, X have singular value decomposition $B = U\Sigma V, X = W\Lambda Z$, where $U, W \in \mathbb{R}^{m \times m}, \Sigma, \Lambda \in \mathbb{R}^{m \times n}, V, Z \in \mathbb{R}^{n \times n}$. Then

$$\begin{aligned} \langle B, X \rangle &= \text{tr}(B^T X) = \text{tr}(V^T \Sigma^T U^T W \Lambda Z) = \text{tr}(\Sigma^T U^T W \Lambda Z V^T) \\ &= \text{diag}(\Sigma) \cdot \text{diag}(U^T W \Lambda Z V^T) \end{aligned}$$

Since the rank of B is at most r , $\text{diag}(\Sigma)$ is supported on the first r entries,

$$\begin{aligned} |\langle B, X \rangle| &\leq \sqrt{\sum_{i=1}^r \Sigma_{ii}^2} \cdot \sqrt{\sum_{i=1}^r (U^T W \Lambda Z V^T)_{ii}^2} \\ &\leq \|B\|_F \sqrt{\sum_{i=1}^r \sum_{j=1}^n (U^T W \Lambda Z V^T)_{ij}^2} = \|B\|_F \|K \Lambda Z V^T\|_F \end{aligned}$$

where we note $K \in \mathbb{R}^{r \times n}$ as the first r rows of $U^T W$. In addition,

$$\|K \Lambda Z V^T\|_F^2 = \text{tr}(V Z^T \Lambda^T K^T K \Lambda Z V^T) = \text{tr}(\Lambda Z V^T V Z^T \Lambda^T K^T K) = \text{tr}(\Lambda^2 K^T K)$$

By K is the first r row of an $n \times n$ orthogonal matrix, we have $\text{tr}(K^T K) = \text{tr}(K K^T) = \text{tr}(I_r) = r$ and all diagonal elements of $K^T K$ are in $[0, 1]$, then

$$\text{tr}(\Lambda^2 K^T K) = \sum_{i=1}^n \lambda_i^2 (K^T K)_{ii} \leq \sum_{i=1}^r \lambda_i^2$$

In summary,

$$|\langle B, X \rangle| \leq \|B\|_F \|K \Lambda Z V^T\|_F \leq \|B\|_F \sqrt{\sum_{i=1}^r \lambda_i^2}. \quad \square$$

It is noteworthy that the signal version of this lemma simply holds by Cauchy-Schwarz inequality.

As seen in the proofs of Theorems 1.1.1 and 1.2.1, it is necessary to estimate the left hand side of (A.9), (A.10), (A.14), (A.15) and (A.16). Notice that these terms are of the similar type – they are all the differences of the squared Frobenius norm of two matrices which only differ on a few leading terms in their SVD decompositions, we have the following lemma for the general estimation of this type of differences. Before we present the lemma, recall that we have defined the concept of indicator vector in the proof of Part 1 in Theorem 1.1.1.

Lemma A.1.4. *For the vector case, suppose $g, h \geq 0, g + h \leq k$,*

$$\{d_i\}_{i=1}^g, \{e_j\}_{j=1}^h, \{t_{ij}\}_{1 \leq i \leq g, 1 \leq j \leq h}$$

are non-negative real numbers satisfying

$$\min_{1 \leq i \leq g} d_i \geq \max_{1 \leq i \leq l} e_i, \quad (\text{A.46})$$

$$\sum_{i=1}^g t_{ij} = e_j, \quad \forall 1 \leq j \leq l \quad (\text{A.47})$$

$\{b_i\}_{i=1}^h, \{c_i\}_{i=1}^h$ are real numbers. $\{u_{11}, \dots, u_{1h}; u_{31}, \dots, u_{3g}; u_{41}, \dots, u_{4l}\}$ is a set of indicator vectors with different support in \mathbb{R}^p ; $\{u_{21}, \dots, u_{2h}; u_{31}, \dots, u_{3g}; u_{41}, \dots, u_{4l}\}$ is also a set of indicator vectors with different support. Define

$$\beta_1 = \sum_{i=1}^h b_i u_{1i} + \sum_{i=1}^g d_i u_{3i} + \sum_{j=1}^l e_j u_{4j} \in \mathbb{R}^p$$

$$\beta_2 = \sum_{i=1}^h c_i u_{2i} + \sum_{i=1}^g d_i u_{3i} + \sum_{j=1}^l e_j u_{4j} \in \mathbb{R}^p$$

Then we have

$$\begin{aligned} \|A\beta_1\|_2^2 - \|A\beta_2\|_2^2 &\geq (1 - \delta_k^A) \left(\sum_{i=1}^h b_i^2 + \sum_{i=1}^g (d_i + \sum_{j=1}^l t_{ij})^2 \right) \\ &\quad - (1 + \delta_k^A) \left(\sum_{i=1}^h c_i^2 + \sum_{i=1}^g (d_i + \sum_{j=1}^l t_{ij})^2 \right) \end{aligned} \quad (\text{A.48})$$

For the matrix case, suppose $g, h \geq 0, g + h \leq r$, $\{d_i\}_{i=1}^g, \{e_j\}_{j=1}^l, \{t_{ij}\}_{1 \leq i \leq g, 1 \leq j \leq l}$ are non-negative real numbers satisfying

$$\min_{1 \leq i \leq g} d_i \geq \max_{1 \leq i \leq l} e_i, \quad (\text{A.49})$$

$$\sum_{i=1}^g t_{ij} = e_j, \quad \forall 1 \leq j \leq l \quad (\text{A.50})$$

$\{b_i\}_{i=1}^h, \{c_i\}_{i=1}^h$ are real numbers. $\{u_{31}, \dots, u_{3g}; u_{41}, \dots, u_{4l}\}$ is a set of orthogo-

nal unit vectors in \mathbb{R}^m , $\{u_{11}, \dots, u_{1h}\}$ and $\{u_{21}, \dots, u_{2h}\}$ are two sets of orthogonal unit vectors lying in the perpendicular space of $\text{span}\{u_{31}, \dots, u_{3g}; u_{41}, \dots, u_{4l}\}$; $\{v_{31}, \dots, v_{3g}; v_{41}, \dots, v_{4l}\}$ is a set of orthogonal unit vectors in \mathbb{R}^n , $\{v_{11}, \dots, v_{1h}\}$ and $\{v_{21}, \dots, v_{2h}\}$ are two sets of orthogonal unit vectors lying in the perpendicular space of $\text{span}\{v_{31}, \dots, v_{3g}; v_{41}, \dots, v_{4l}\}$. Define

$$X_1 = \sum_{i=1}^h b_i u_{1i} v_{1i}^T + \sum_{i=1}^g d_i u_{3i} v_{3i}^T + \sum_{j=1}^l e_j u_{4j} v_{4j}^T \in \mathbb{R}^{m \times n}$$

$$X_2 = \sum_{i=1}^h c_i u_{2i} v_{2i}^T + \sum_{i=1}^g d_i u_{3i} v_{3i}^T + \sum_{j=1}^l e_j u_{4j} v_{4j}^T \in \mathbb{R}^{m \times n}$$

Then we have

$$\begin{aligned} \|\mathcal{M}(X_1)\|_2^2 - \|\mathcal{M}(X_2)\|_2^2 &\geq (1 - \delta_r^{\mathcal{M}}) \left(\sum_{i=1}^h b_i^2 + \sum_{i=1}^g (d_i + \sum_{j=1}^l t_{ij})^2 \right) \\ &\quad - (1 + \delta_r^{\mathcal{M}}) \left(\sum_{i=1}^h c_i^2 + \sum_{i=1}^g (d_i + \sum_{j=1}^l t_{ij})^2 \right) \end{aligned} \quad (\text{A.51})$$

Proof of Lemma A.1.4. We prove the Lemma by induction on l . We prove matrix case only as the signal case is essentially the same.

When $l = 0$, (A.51) is clear to hold by the definition of $\delta_r^{\mathcal{M}}$ and the fact that $g + h \leq r$. Suppose (A.51) holds for $l - 1$, ($l \geq 1$), we note

$$Y_i = -u_{3i} v_{3i}^T + u_{4i} v_{4i}^T, \quad 1 \leq i \leq g \quad (\text{A.52})$$

$$P_z = X_z - \sum_{i=1}^g t_{il} Y_i, \quad z = 1, 2 \quad (\text{A.53})$$

$$Q_{iz} = X_z - \sum_{w=1}^g t_{wl} Y_w + (t_{il} + d_i) Y_i \quad z = 1, 2, \quad 1 \leq i \leq g \quad (\text{A.54})$$

We can show the following equality in l_2 -space:

$$\begin{aligned} & \mu \left\| \mathcal{M}(X_z - \sum_{i=1}^g t_{il} Y_i) \right\|_2^2 + \sum_{i=1}^g \nu_i \left\| \mathcal{M}(X_z - \sum_{w=1}^g t_{wl} Y_w + (t_{il} + d_i) Y_i) \right\|_2^2 \\ &= \left\| \mathcal{M}(X_z) \right\|_2^2 + \mu \left\| \mathcal{M}(\sum_{i=1}^g t_{il} Y_i) \right\|_2^2 + \sum_{i=1}^g \nu_i \left\| \mathcal{M}(-\sum_{w=1}^g t_{wl} Y_w + (t_{il} + d_i) Y_i) \right\|_2^2 \end{aligned} \quad (\text{A.55})$$

where $z = 1, 2$, $\nu_i = \frac{t_{il}}{d_i + t_{il}}$, $\mu = 1 - \sum_{i=1}^g \frac{t_{il}}{d_i + t_{il}}$. By (A.49), (A.50) we have

$$\mu \geq 1 - \sum_{i=1}^g \frac{t_{il}}{d_i} = 1 - \frac{e_l}{d_i} \geq 0$$

Thus, ν_i, μ are all non-negative numbers satisfying $\mu + \sum_{i=1}^g \nu_i = 1$. Consider the difference of these two equalities (A.55) ($z = 1, 2$), we get

$$\begin{aligned} & \left\| \mathcal{M}(X_1) \right\|_2^2 - \left\| \mathcal{M}(X_2) \right\|_2^2 \\ &= \mu \left[\left\| \mathcal{M}(P_1) \right\|_2^2 - \left\| \mathcal{M}(P_2) \right\|_2^2 \right] + \sum_{i=1}^g \nu_i \left[\left\| \mathcal{M}(Q_{i1}) \right\|_2^2 - \left\| \mathcal{M}(Q_{i2}) \right\|_2^2 \right] \end{aligned} \quad (\text{A.56})$$

By computing directly we can get

$$\begin{aligned} P_1 &= \sum_{i=1}^h b_i u_{1i} v_{1i}^T + \sum_{i=1}^g (d_i + t_{il}) u_{3i} v_{3i}^T + \sum_{j=1}^{l-1} e_j u_{4j} v_{4j}^T \\ P_2 &= \sum_{i=1}^h c_i u_{2i} v_{2i}^T + \sum_{i=1}^g (d_i + t_{il}) u_{3i} v_{3i}^T + \sum_{j=1}^{l-1} e_j u_{4j} v_{4j}^T \\ Q_{i1} &= \sum_{w=1}^h b_w u_{1w} v_{1w}^T + \left[\sum_{w=1, w \neq i}^g (d_w + t_{wl}) u_{3w} v_{3w}^T + (d_i + t_{il}) u_{4l} v_{4l}^T \right] + \sum_{j=1}^{l-1} e_j u_{4j} v_{4j}^T \\ Q_{i2} &= \sum_{w=1}^h c_w u_{2w} v_{2w}^T + \left[\sum_{w=1, w \neq i}^g (d_w + t_{wl}) u_{3w} v_{3w}^T + (d_i + t_{il}) u_{4l} v_{4l}^T \right] + \sum_{j=1}^{l-1} e_j u_{4j} v_{4j}^T \end{aligned}$$

which corresponds with the assumption of $l - 1$. Now by induction assumption of

$l - 1$, for all $1 \leq w \leq g$ we have

$$\begin{aligned}
& \|\mathcal{M}(P_1)\|_2^2 - \|\mathcal{M}(P_2)\|_2^2 \geq (1 - \delta_r^{\mathcal{M}}) \left(\sum_{i=1}^h b_i^2 + \sum_{i=1}^g (d_i + \sum_{j=1}^l t_{ij})^2 \right) \\
& - (1 + \delta_r^{\mathcal{M}}) \left(\sum_{i=1}^h c_i^2 + \sum_{i=1}^g (d_i + \sum_{j=1}^l t_{ij})^2 \right) \\
& \|\mathcal{M}(Q_{w1})\|_2^2 - \|\mathcal{M}(Q_{w2})\|_2^2 \geq (1 - \delta_r^{\mathcal{M}}) \left(\sum_{i=1}^h b_i^2 + \sum_{i=1}^g (d_i + \sum_{j=1}^l t_{ij})^2 \right) \\
& - (1 + \delta_r^{\mathcal{M}}) \left(\sum_{i=1}^h c_i^2 + \sum_{i=1}^g (d_i + \sum_{j=1}^l t_{ij})^2 \right)
\end{aligned} \tag{A.57}$$

Together (A.57) with (A.56), we can get (A.51) for the case l . \square

Lemma A.1.5. *Suppose $m \geq r$, $a_1 \geq a_2 \geq \dots \geq a_m \geq 0$, $\sum_{i=1}^r a_i \geq \sum_{i=r+1}^m a_i$, then for all $\alpha \geq 1$,*

$$\sum_{j=r+1}^m a_j^\alpha \leq \sum_{i=1}^r a_i^\alpha. \tag{A.58}$$

More generally, suppose $a_1 \geq a_2 \geq \dots \geq a_m \geq 0$, $\lambda \geq 0$ and $\sum_{i=1}^r a_i + \lambda \geq \sum_{i=r+1}^m a_i$, then for all $\alpha \geq 1$,

$$\sum_{j=r+1}^m a_j^\alpha \leq r \left(\sqrt[\alpha]{\frac{\sum_{i=1}^r a_i^\alpha}{r}} + \frac{\lambda}{r} \right)^\alpha \tag{A.59}$$

Proof of Lemma A.1.5. It is sufficient to show the general part only. Since we can set $a_j = 0$ when $j > m$, we assume $m \geq 2r$ without loss of generality. By Lemma A.1.1, we can find $\{s_{ij}\}_{1 \leq i \leq r, 2r+1 \leq j \leq m}$ satisfying (A.25), (A.26). Hence,

$$\begin{aligned}
\sum_{j=r+1}^m a_j^\alpha &= \sum_{j=2r+1}^m a_j^{\alpha-1} \left(\sum_{i=1}^r s_{ij} \right) + \sum_{j=r+1}^{2r} a_j^\alpha = \sum_{i=1}^r \left(a_{r+i}^\alpha + \sum_{j=2r+1}^m a_j^{\alpha-1} s_{ij} \right) \\
&\leq \sum_{i=1}^r a_{r+i}^{\alpha-1} \left(a_{r+i} + \sum_{j=2r+1}^m s_{ij} \right) \leq \sum_{i=1}^r \left(a_{r+i} + \sum_{j=2r+1}^m s_{ij} \right)^\alpha \\
&\leq r \left(\frac{\sum_{i=1}^r a_i}{r} + \frac{\lambda}{r} \right)^\alpha \leq r \left(\sqrt[\alpha]{\frac{\sum_{i=1}^r a_i^\alpha}{r}} + \frac{\lambda}{r} \right)^\alpha. \quad \square
\end{aligned}$$

It is also interesting to consider conditions on δ_{sk}^A and δ_{sr}^M for some real number $s > 1$. The following result provides convenient bounds on δ_{sk}^A and δ_{sr}^M in terms of δ_k^A and δ_r^M respectively. It is also useful for the proof of Proposition 1.2.2.

Lemma A.1.6. *For all matrix $A \in \mathbb{R}^{n \times p}$ and $k \geq 2$ is an integer, $s > 1$ is real and sk is integer. Then we have $\delta_{sk}^A \leq (2s - 1)\delta_k^A$. Similarly, for all linear map $\mathcal{M} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^q$ and $r \geq 2$, $s > 2$ and sr is integer. Then we have $\delta_{sr}^M \leq (2s - 1)\delta_r^M$.*

Proof of Lemma A.1.6 We only show the matrix case. For all $X \in \mathbb{R}^{m \times n}$ such that $\text{rank}(X) \leq sr$, suppose X has singular value decomposition $X = \sum_{i=1}^l a_i u_i v_i^T$, $l \leq sr$. Without loss of generality we can assume $l = sr$ as we can set $a_i = 0$ if $l < i \leq sr$. Note

$$w_i = \mathcal{M}(a_i u_i v_i^T) \in \mathbb{R}^q, \quad 1 \leq i \leq sr$$

We can verify the following identity

$$\begin{aligned} & \left\| \sum_{i=1}^{sr} w_i \right\|_2^2 + \frac{s-1}{sr-1} \sum_{1 \leq i < j \leq sr} \|w_i - w_j\|_2^2 \\ &= (1 + (s-1)) \sum_{i=1}^{sr} \|w_i\|_2^2 + 2\left(1 - \frac{s-1}{sr-1}\right) \sum_{1 \leq i < j \leq sr} \langle w_i, w_j \rangle \\ &= \frac{s^2}{\binom{sr}{r}} \sum_{1 \leq i_1 < \dots < i_r \leq sr} \|w_{i_1} + w_{i_2} + \dots + w_{i_r}\|_2^2 \end{aligned}$$

which implies

$$\begin{aligned}
\|\mathcal{M}(X)\|_2^2 &= \left\| \sum_{i=1}^{sr} w_i \right\|_2^2 \\
&\leq \frac{s^2(1 + \delta_r^{\mathcal{M}})}{\binom{sr}{r}} \sum_{1 \leq i_1 < \dots < i_r \leq sr} (a_{i_1}^2 + \dots + a_{i_r}^2) \\
&\quad - \frac{(s-1)(1 - \delta_r^{\mathcal{M}})}{sr-1} \sum_{1 \leq i < j \leq sr} (a_i^2 + a_j^2) \\
&= (s(1 + \delta_r^{\mathcal{M}}) - (s-1)(1 - \delta_r^{\mathcal{M}})) \sum_{i=1}^{rs} a_i^2 \\
&= (1 + (2s-1)\delta_r^{\mathcal{M}}) \|X\|_F^2
\end{aligned}$$

$$\begin{aligned}
\|\mathcal{M}(X)\|_2^2 &= \left\| \sum_{i=1}^{sr} w_i \right\|_2^2 \\
&\geq \frac{s^2(1 - \delta_r^{\mathcal{M}})}{\binom{sr}{r}} \sum_{1 \leq i_1 < \dots < i_r \leq sr} (a_{i_1}^2 + \dots + a_{i_r}^2) \\
&\quad - \frac{(s-1)(1 + \delta_r^{\mathcal{M}})}{sr-1} \sum_{1 \leq i < j \leq sr} (a_i^2 + a_j^2) \\
&= (s(1 - \delta_r^{\mathcal{M}}) - (s-1)(1 + \delta_r^{\mathcal{M}})) \sum_{i=1}^{rs} a_i^2 \\
&= (1 - (2s-1)\delta_r^{\mathcal{M}}) \|X\|_F^2
\end{aligned}$$

Hence, $\delta_{sr}^{\mathcal{M}} \leq (2s-1)\delta_r^{\mathcal{M}}$. \square

Lemma A.1.7, which reveals the relationship between ROC's of different orders, is from Cai et al. (2010c).

Lemma A.1.7. *For any $\mu \geq 1$ and positive integers k_1, k_2 such that μk_2 is an integer, then*

$$\theta_{k_1, \mu k_2} \leq \sqrt{\mu} \theta_{k_1, k_2}$$

The following Lemma A.1.8 provides a bound for the ROC θ in terms of the RIC

δ and can be used to compare different RIP conditions.

Lemma A.1.8. *Let $A \in \mathbb{R}^{n \times p}$. Then we have*

$$\theta_{k,k}^A \leq \begin{cases} 2\delta_k^A, & \text{when } k \text{ is even, } k \geq 2; \\ \frac{2k}{\sqrt{k^2-1}}\delta_k^A, & \text{when } k \text{ is odd, } k \geq 3. \end{cases} \quad (\text{A.60})$$

In addition, both coefficients, 2 in the even case and $\frac{2k}{\sqrt{k^2-1}}$ in the odd case, cannot be further improved.

Similarly, in the matrix case, for a linear map $\mathcal{M} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^q$,

$$\theta_{r,r}^{\mathcal{M}} \leq \begin{cases} 2\delta_r^{\mathcal{M}}, & \text{when } r \text{ is even, } r \geq 2; \\ \frac{2r}{\sqrt{r^2-1}}\delta_r^{\mathcal{M}}, & \text{when } r \text{ is odd, } r \geq 3. \end{cases} \quad (\text{A.61})$$

In addition, the coefficient 2 in the even case cannot be further improved.

Proof of Lemma A.1.8. For k -sparse vectors $\beta, \gamma \in \mathbb{R}^p$ with disjoint supports, we can write them as $\beta = \sum_{i \in T_1} a_i e_i$ and $\gamma = \sum_{i \in T_2} b_i e_i$ where $a_i > 0$, $b_i > 0$, T_1 is the support of β , T_2 is the support of γ , and e_i is the vector with i th entry equals to ± 1 and all other entries equal to zero. Correspondingly, suppose $X, Y \in \mathbb{R}^{m \times n}$ with rank at most r , which satisfies $X^T Y = X Y^T = 0$. Lemma 2.3 in Recht et al. (2010) shows that they have singular value decompositions $X = \sum_{i \in T_1} a_i u_i v_i^T$ and $Y = \sum_{i \in T_2} b_i u_i v_i^T$, where the disjoint subsets T_1 and T_2 satisfy $|T_1|, |T_2| \leq r$. We now consider the even and odd cases separately.

Case 1. $k, r \geq 2$ is even. We focus on the matrix case. The proof of the signal case is similar. Without loss of generality, suppose X and Y are normalized so $\|X\|_F = \|Y\|_F = 1$. Divide T_1 and T_2 into two parts, $T_1 = T_{11} \cup T_{12}$, $T_2 = T_{21} \cup T_{22}$, such that $T_{11}, T_{12}, T_{21}, T_{22}$ are disjoint and $|T_{ij}| \leq r/2$ for $i, j \in \{1, 2\}$. Denote

$X_i = \sum_{i \in T_{1i}} a_i u_i v_i^T$ and $Y_i = \sum_{i \in T_{2i}} b_i u_i v_i^T$, $i = 1, 2$. Then

$$\begin{aligned}
|\langle \mathcal{M}(X), \mathcal{M}(Y) \rangle| &\leq \sum_{i,j=1}^2 |\langle \mathcal{M}(X_i), \mathcal{M}(Y_j) \rangle| \\
&= \frac{1}{4} \sum_{i,j=1}^2 \left| \|\mathcal{M}(X_i + Y_j)\|_F^2 - \|\mathcal{M}(X_i - Y_j)\|_F^2 \right| \\
&\leq \frac{1}{4} \sum_{i,j=1}^2 \left[(1 + \delta_r^{\mathcal{M}}) \sum_{i \in T_{ij} \cup T_{ij}} a_i^2 - (1 - \delta_r^{\mathcal{M}}) \sum_{i \in T_{ij} \cup T_{ij}} a_i^2 \right] \\
&= \delta_r^{\mathcal{M}} (\|X\|_F^2 + \|Y\|_F^2) = 2\delta_r^{\mathcal{M}}
\end{aligned}$$

and consequently $\theta_{r,r}^{\mathcal{M}} \leq 2\delta_r^{\mathcal{M}}$. Now in the example provided in the proof of Theorem 1.2.4, if $a = b = k$, we have $\delta_r^A = 1/3$, $\theta_{r,r}^{\mathcal{M}} = 2/3$, which means the coefficient “2” in the inequalities of the even case in (A.61) cannot be improved.

Case 2. $k, r \geq 3$ is odd. For the proof of (A.60) and (A.61), we only show the matrix case as the signal case is similar. Since we can set $a_i = 0$ or $b_i = 0$ for $i \notin T_1$ or $i \notin T_2$, Without loss of generality, we assume that $|T_1| = r, |T_2| = r$, a_i, b_i might be 0 for $i \in T_1 \cup T_2$. Also without loss of generality, we can assume X and Y are

normalized so $\|X\|_F^2 = \sum_{i \in T_1} a_i^2 = \sqrt{\frac{r-1}{r+1}}$ and $\|Y\|_F^2 = \sum_{i \in T_2} b_i^2 = \sqrt{\frac{r+1}{r-1}}$. Then

$$\begin{aligned}
& \left| 4 \binom{r-1}{(r-1)/2} \binom{r-1}{(r-3)/2} \langle \mathcal{M}(X), \mathcal{M}(Y) \rangle \right| \\
&= \left| 4 \binom{r-1}{(r-1)/2} \binom{r-1}{(r-3)/2} \langle \mathcal{M}(\sum_{i \in T_1} a_i u_i v_i^T), \mathcal{M}(\sum_{i \in T_2} b_i u_i v_i^T) \rangle \right| \\
&= \left| \sum_{\substack{A \subseteq T_1, |A|=(r+1)/2, \\ B \subseteq T_2, |B|=(r-1)/2}} \left[\|\mathcal{M}(\sum_{i \in A} a_i u_i v_i^T + \sum_{i \in B} b_i u_i v_i^T)\|^2 \right. \right. \\
&\quad \left. \left. - \|\mathcal{M}(\sum_{i \in A} a_i u_i v_i^T - \sum_{i \in B} b_i u_i v_i^T)\|^2 \right] \right| \\
&\leq \sum_{\substack{A \subseteq T_1, |A|=(r+1)/2, \\ B \subseteq T_2, |B|=(r-1)/2}} ((1 + \delta_r^{\mathcal{M}}) - (1 - \delta_r^{\mathcal{M}})) \left[\sum_{i \in A} a_i^2 + \sum_{i \in B} b_i^2 \right] \\
&= 2\delta_r^{\mathcal{M}} \left[\binom{r-1}{(r-1)/2} \binom{r}{(r-1)/2} \sum_{i \in T_1} a_i^2 + \binom{r-1}{(r-3)/2} \binom{r}{(r+1)/2} \sum_{i \in T_2} b_i^2 \right] \\
&= 2\delta_r^{\mathcal{M}} \binom{r-1}{(r-1)/2} \binom{r-1}{(r-3)/2} \left[\frac{r}{(r-1)/2} \sum_{i \in T_1} a_i^2 + \frac{r}{(r+1)/2} \sum_{i \in T_2} b_i^2 \right] \\
&= 8\delta_r^{\mathcal{M}} \binom{r-1}{(r-1)/2} \binom{r-1}{(r-3)/2} \frac{r}{\sqrt{r^2-1}} \\
&= 4 \binom{r-1}{(r-1)/2} \binom{r-1}{(r-3)/2} \frac{2r}{\sqrt{r^2-1}} \delta_r^{\mathcal{M}} \|X\|_F \|Y\|_F
\end{aligned}$$

which implies $\theta_{r,r}^{\mathcal{M}} \leq \frac{2r}{\sqrt{r^2-1}} \delta_r^{\mathcal{M}}$.

Next we will construct an example for the signal recovery in the odd case where

$\theta_{k,k}^A = \frac{2k}{\sqrt{k^2-1}} \delta_k^A \neq 0$. Suppose $k \geq 3$ is odd and $2k \leq p$, denote

$$\beta_1 = \frac{1}{\sqrt{2k}} (\overbrace{1, 1, \dots, 1}^{2k}, 0, \dots) \in \mathbb{R}^p \quad \beta_2 = \frac{1}{\sqrt{2k}} (\overbrace{1, 1, \dots, 1}^k, \overbrace{-1, \dots, -1}^k, 0, \dots) \in \mathbb{R}^p. \tag{A.62}$$

Similarly as in the proof of Theorem 1.2.4, we can extend β_1 and β_2 to an orthonormal

basis of \mathbb{R}^p as $\{\beta_1, \beta_2, \dots, \beta_p\}$. Then for $0 < \lambda < 1$, we define $A : \mathbb{R}^p \rightarrow \mathbb{R}^p$ by

$$A\beta = \sqrt{1 + \lambda}a_1\beta_1 + \sqrt{1 - \lambda}a_2\beta_2 + \sum_{i=3}^p a_i\beta_i$$

for $\beta = \sum_{i=1}^p a_i\beta_i$. It is clear that for all $\beta \in \mathbb{R}^p$, $(1 - \lambda)\|\beta\|_2^2 \leq \|A\beta\|_2^2 \leq (1 + \lambda)\|\beta\|_2^2$.

Let β and γ be k -sparse vectors with disjoint supports and $\|\beta\|_2 = \|\gamma\|_2 = 1$. Then

$$\begin{aligned} & |\langle A\beta, A\gamma \rangle| \\ &= \frac{1}{4} \left| \|A(\beta + \gamma)\|_2^2 - \|A(\beta - \gamma)\|_2^2 \right| \\ &\leq \max \left\{ \frac{1 + \lambda}{4} \|\beta + \gamma\|_2^2 - \frac{1 - \lambda}{4} \|\beta - \gamma\|_2^2, \frac{1 + \lambda}{4} \|\beta - \gamma\|_2^2 - \frac{1 - \lambda}{4} \|\beta + \gamma\|_2^2 \right\} \\ &= \frac{2\lambda}{4} (\|\beta\|_2^2 + \|\gamma\|_2^2) = \lambda \|\beta\|_2 \|\gamma\|_2 \end{aligned}$$

which implies $\theta_{k,k}^A \leq \lambda$. It can be easily verified that when

$$\beta = (\overbrace{1, 1, \dots, 1}^k, 0, \dots) \quad \text{and} \quad \gamma = (\overbrace{0, 0, \dots, 0}^k, \overbrace{1, 1, \dots, 1}^k, 0, \dots),$$

we have $|\langle A\beta, A\gamma \rangle| = \lambda \|\beta\|_2 \|\gamma\|_2$. These together imply $\theta_{k,k}^A = \lambda$.

Denote $\beta(i)$ as the i th entry of β . Now let us estimate δ_k^A . For all k -sparse $\beta \in \mathbb{R}^p$, suppose $\beta = \sum_{i=1}^p c_i\beta_i$, then

$$\begin{aligned} \|A\beta\|_2^2 &= (1 + \lambda)|\langle \beta, \beta_1 \rangle|^2 + (1 - \lambda)|\langle \beta, \beta_2 \rangle|^2 + \sum_{i=3}^p |\langle \beta, \beta_i \rangle|^2 \\ &= \|\beta\|_2^2 + \lambda(|\langle \beta, \beta_1 \rangle|^2 - |\langle \beta, \beta_2 \rangle|^2) \\ &= \|\beta\|_2^2 + \lambda \left(\left(\sum_{i=1}^{2k} \beta(i) \right)^2 - \left(\sum_{i=1}^k \beta(i) - \sum_{i=k+1}^{2k} \beta(i) \right)^2 \right) / 2k \\ &= \|\beta\|_2^2 + \frac{4}{2k} \lambda \left(\sum_{i=1}^k \beta(i) \right) \left(\sum_{i=j+1}^{2k} \beta(i) \right). \end{aligned}$$

Suppose $T_1 = \text{supp}(\beta) \cap \{1, \dots, k\}$ and $T_2 = \text{supp}(\beta) \cap \{k+1, \dots, 2k\}$, then $|T_1| + |T_2| \leq k$ and

$$\begin{aligned}
& \left| \left(\sum_{i=1}^k \beta(i) \right) \left(\sum_{i=k+1}^{2k} \beta(i) \right) \right| \\
&= \left| \left(\sum_{i \in T_1} \beta(i) \right) \left(\sum_{i \in T_2} \beta(i) \right) \right| \leq \sqrt{|T_1| \sum_{i \in T_1} \beta(i)^2 \cdot |T_2| \sum_{i \in T_2} \beta(i)^2} \\
&\leq \frac{\sqrt{|T_1| \cdot |T_2|}}{2} \sum_{i \in T_1 \cup T_2} \beta(i)^2 \leq \frac{\sqrt{|T_1|(k - |T_1|)}}{2} \|\beta\|_2^2 \leq \frac{\sqrt{\frac{k-1}{2} \frac{k+1}{2}}}{2} \|\beta\|_2^2,
\end{aligned}$$

where the last inequality is due to the facts that $|T_1|$ is a non-negative integer and k is odd. It then follows that for all k -sparse vector $\beta \in \mathbb{R}^p$,

$$\left(1 - \frac{\sqrt{k^2 - 1}}{2k} \lambda\right) \|\beta\|_2^2 \leq \|A\beta\|_2^2 \leq \left(1 + \frac{\sqrt{k^2 - 1}}{2k} \lambda\right) \|\beta\|_2^2.$$

It can also be easily verified that the equality above can be achieved for

$$\beta = \left(\overbrace{1, \dots, 1}^{(k+1)/2}, \overbrace{0, \dots, 0}^{(k-1)/2}, \overbrace{1, \dots, 1}^{(k-1)/2}, 0, \dots \right)$$

Hence $\delta_k^A = \lambda \frac{\sqrt{k^2 - 1}}{2k}$. In summary, $\theta_{k,k}^A = \frac{2k}{\sqrt{k^2 - 1}} \delta_k^A$ in our setting, which implies that the constant $\frac{2k}{\sqrt{k^2 - 1}}$ in (A.60) is not improvable. \square

A.2 Supplement for Chapter 2

We prove the main results of Chapter 2 in this Appendix. We begin by collecting a few important technical lemmas that will be used in the proofs of the main results. The proofs of some of these technical lemmas are involved and are postponed to Section A.2.11.

A.2.1 Technical Tools

Lemmas A.2.1 and A.2.2 below are used for deriving the RUB condition (see Definition 2.2.1) from the ROP design.

Lemma A.2.1. *Suppose $A \in \mathbb{R}^{p_1 \times p_2}$ is a fixed matrix and \mathcal{X} is ROP from a symmetric sub-Gaussian distribution \mathcal{P} , i.e.*

$$[\mathcal{X}(A)]_j = \beta^{(j)T} A \gamma^{(j)}, \quad j = 1, \dots, n$$

where $\beta^{(j)} = (\beta_1^{(j)}, \dots, \beta_{p_1}^{(j)})^\top, \gamma^{(j)} = (\gamma_1^{(j)}, \dots, \gamma_{p_2}^{(j)})^\top$ are random vectors with entries i.i.d. generated from \mathcal{P} . Then for $\delta > 0$, we have

$$\left(\frac{1}{3\alpha_{\mathcal{P}}^4} - 2\alpha_{\mathcal{P}}^2\delta - \alpha_{\mathcal{P}}^2\delta^2 \right) \|A\|_F \leq \|\mathcal{X}(A)\|_1/n \leq (1 + 2\alpha_{\mathcal{P}}^2\delta + \alpha_{\mathcal{P}}^2\delta^2) \|A\|_F$$

with probability at least $1 - 2\exp(-\delta^2 n)$. Here $\alpha_{\mathcal{P}}$ is defined by (2.26).

Lemma A.2.2. *Suppose $A \in \mathbb{R}^{p_1 \times p_2}$ is a fixed matrix. $\beta = (\beta_1, \dots, \beta_{p_1})^\top, \gamma = (\gamma_1, \dots, \gamma_{p_2})^\top$ are random vectors such that $\beta_1, \dots, \beta_{p_1}, \gamma_1, \dots, \gamma_{p_2} \stackrel{iid}{\sim} \mathcal{P}$, where \mathcal{P} is some symmetric variance 1 sub-Gaussian distribution, then we have*

$$\frac{\|A\|_F}{3\alpha_{\mathcal{P}}^4} \leq E |\beta^\top A \gamma| \leq \|A\|_F$$

where $\alpha_{\mathcal{P}}$ is given by (2.26).

Let $z \in \mathbb{R}^n$ be i.i.d sub-Gaussian distributed. By measure concentration theory, $\|z\|_p^p/n$, $1 \leq p \leq \infty$, are essentially bounded; Specifically, we have the following lemma.

Lemma A.2.3. *Suppose $z \in \mathbb{R}^n$ and $z_i \stackrel{iid}{\sim} N(0, \sigma^2)$, we have*

$$\begin{aligned} P(\|z\|_1 \geq \sigma n) &\leq \frac{9}{n} \\ P\left(\|z\|_2 \geq \sigma \sqrt{n + 2\sqrt{n \log n}}\right) &\leq \frac{1}{n} \\ P(\|z\|_\infty \geq 2\sigma \sqrt{\log n}) &\leq \frac{1}{n\sqrt{2\pi \log n}}. \end{aligned}$$

More general, when z_i are i.i.d. sub-Gaussian distributed such that (2.25) holds, then

$$\begin{aligned} P(\|z\|_1 \geq Cn) &\leq \exp\left(-\frac{n(C - 2\sqrt{2\pi}\gamma)^2}{2\gamma^2}\right), \quad \forall C > 2\sqrt{2\pi}\gamma \\ P(\|z\|_2 \geq \sqrt{Cn}) &\leq \exp\left(-\frac{n(C - 4\gamma^2)^2}{8\gamma^2 C}\right), \quad \forall C > 4\gamma^2 \\ P(\|z\|_\infty \geq C\gamma\sqrt{\log n}) &\leq 2n^{-C^2/2-1}, \quad \forall C > 0 \end{aligned}$$

Lemma A.2.4 below presents an upper bound for the spectral norm of $\mathcal{X}(z)$ for a fixed vector z .

Lemma A.2.4. *Suppose \mathcal{X} is ROP from some symmetric sub-Gaussian distribution \mathcal{P} and $z \in \mathbb{R}^n$ is some fixed vector, then for $C > \log 7$, we have*

$$\|\mathcal{X}^*(z)\| \leq 3\alpha_{\mathcal{P}}^2 \left(C(p_1 + p_2)\|z\|_\infty + \sqrt{2C(p_1 + p_2)}\|z\|_2 \right)$$

with probability at least $1 - 2 \exp(-(C - \log 7)(p_1 + p_2))$. Here $\alpha_{\mathcal{P}}$ is defined by (2.26).

We are now ready to prove the main results of Chapter 2.

A.2.2 Proof of Theorem 2.2.1.

For the proof of Theorem 2.2.1, by null space property (Lemma A.1.2), we only need to show for all non-zero R with $\mathcal{X}(R) = 0$, we must have $\|R_{\max(r)}\|_* < \|R_{-\max(r)}\|_*$.

If this does not hold, suppose there exists non-zero R with $\mathcal{X}(R) = 0$ and $\|R_{\max(r)}\|_* \geq \|R_{-\max(r)}\|_*$. We denote $p = \min(p_1, p_2)$ and assume the singular value decomposition of R is

$$R = \sum_{i=1}^p \sigma_i u_i v_i^\top = U \text{diag}(\vec{\sigma}) V^\top,$$

where u_i, v_i are orthogonal basis in $\mathbb{R}^{p_1}, \mathbb{R}^{p_2}$, respectively and $\vec{\sigma}$ is the singular value vector such that $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$. Without loss of generality, we can assume $p \geq kr$, otherwise we can set the undefined entries of σ as 0.

Consider the singular value vector $\vec{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_p)$, we note that $\vec{\sigma}_{-\max(kr)}$ satisfies

$$\|\vec{\sigma}_{-\max(kr)}\|_\infty \leq \sigma_{kr},$$

$$\begin{aligned} \|\vec{\sigma}_{-\max(kr)}\|_1 &= \|\vec{\sigma}_{-\max(r)}\|_1 - (\sigma_{r+1} + \dots + \sigma_{kr}) \\ &\leq \|\vec{\sigma}_{-\max(r)}\|_1 - (k-1)r\sigma_{kr} \leq \|\vec{\sigma}_{\max(r)}\|_1 - (k-1)r\sigma_{kr}. \end{aligned}$$

Denote $\theta = \max\{\sigma_{kr}, (\|\vec{\sigma}_{\max(r)}\|_1 - r(k-1)\sigma_{kr})/(kr)\}$, by the two inequalities above we have $\|\vec{\sigma}_{-\max(kr)}\|_\infty \leq \theta$ and $\|\vec{\sigma}_{-\max(kr)}\|_1 \leq kr\theta$. Now apply Lemma 1.1.1 in Chapter 1, we can get $b^{(i)} \in \mathbb{R}^p, \lambda_i \geq 0, i = 1, \dots, N$ such that $\sum_{i=1}^N \lambda_i = 1$, $\vec{\sigma}_{-\max(kr)} = \sum_{i=1}^N \lambda_i b^{(i)}$ and

$$\begin{aligned} \text{supp}(b^{(i)}) &\subseteq \text{supp}(\vec{\sigma}_{-\max(kr)}), \quad \|b^{(i)}\|_0 \leq kr, \\ \|b^{(i)}\|_1 &= \|\vec{\sigma}_{-\max(kr)}\|_1, \quad \|b^{(i)}\|_\infty \leq \theta. \end{aligned} \tag{A.63}$$

which leads to

$$\|b^{(i)}\|_2 \leq \sqrt{\|b^{(i)}\|_1 \cdot \|b^{(i)}\|_\infty} \leq \sqrt{(\|\vec{\sigma}_{\max(r)}\|_1 - r(k-1)\sigma_{kr}) \cdot \theta}$$

If $\theta = \sigma_{kr}$, we have

$$\begin{aligned} \|b^{(i)}\|_2 &\leq \sqrt{(\|\vec{\sigma}_{\max(r)}\|_1 - r(k-1)\sigma_{kr})\sigma_{kr}} \\ &\leq \sqrt{\left(\|\vec{\sigma}_{\max(r)}\|_1 - r(k-1)\frac{\|\vec{\sigma}_{\max(r)}\|_1}{2r(k-1)}\right)\frac{\|\vec{\sigma}_{\max(r)}\|_1}{2r(k-1)}} \\ &\leq \frac{\|\vec{\sigma}_{\max(r)}\|_1}{\sqrt{4r(k-1)}} \leq \frac{\|\vec{\sigma}_{\max(r)}\|_2}{\sqrt{4(k-1)}} \end{aligned}$$

If $\theta = (\|\vec{\sigma}_{\max(r)}\|_1 - r(k-1)\sigma_{kr})/(kr)$, we have

$$\|b^{(i)}\|_2 \leq \sqrt{\frac{1}{kr}(\|\vec{\sigma}_{\max(r)}\|_1 - r(k-1)\sigma_{kr})} \leq \sqrt{\frac{1}{kr}}\|\vec{\sigma}_{\max(r)}\|_1 \leq \frac{\|\vec{\sigma}_{\max(r)}\|_2}{\sqrt{k}}$$

Since $k \geq 2$, we always have $\|b^{(i)}\|_2 \leq \|\vec{\sigma}_{\max(r)}\|_2/\sqrt{k}$. Finally, we define $B_i = U\text{diag}(b^{(i)})V^\top$, then the rank of B_i are all at most kr and $\sum_{i=1}^N \lambda_i B_i = R_{-\max(kr)}$ and

$$\|B_i\|_F = \|b^{(i)}\|_2 \leq \|\vec{\sigma}_{\max(r)}\|_2/\sqrt{k} = \|R_{\max(r)}\|_F/\sqrt{k}$$

Hence,

$$\begin{aligned} 0 &= \|\mathcal{X}(R)\|_1 \geq \|\mathcal{X}(R_{\max(kr)})\|_1 - \|\mathcal{X}(R_{-\max(kr)})\|_1 \\ &\geq C_1 \|R_{\max(kr)}\|_F - \sum_{i=1}^N \|\mathcal{X}(\lambda_i B_i)\|_1 \\ &\geq C_1 \|R_{\max(r)}\|_F - \sum_{i=1}^N \lambda_i C_2 \|B_i\|_F \\ &\geq C_1 \|R_{\max(r)}\|_F - C_2 \|R_{\max(r)}\|_F/\sqrt{k} > 0 \end{aligned}$$

Here we used the RUB condition. The last inequality is due to $C_2/C_1 < \sqrt{k}$ and $R \neq 0$ (so $R_{\max(r)} \neq 0$). This is a contradiction, which finishes the proof of the theorem.

A.2.3 Proof of Theorem 2.2.2.

Notice that for \mathcal{P} as standard Gaussian distribution, the constant $\alpha_{\mathcal{P}}$ (defined as (2.26)) equals 1. We will prove the following more general result than Theorem 2.2.2 instead.

Proposition A.2.1. *Suppose $\mathcal{X} : \mathbb{R}^{p_1 \times p_2} \rightarrow \mathbb{R}^n$ is ROP from some variance 1 symmetric sub-Gaussian distribution \mathcal{P} . For integer $k \geq 2$, positive $C_1 < \frac{1}{3\alpha_{\mathcal{P}}^4}$ ($\alpha_{\mathcal{P}}$ is defined as (2.26)) and $C_2 > 1$, there exists constants C and δ , only depending on \mathcal{P}, C_1, C_2 but not on p_1, p_2, r , such that if $n \geq Cr(p_1 + p_2)$, then with probability at least $1 - e^{-n\delta}$, \mathcal{X} satisfies RUB of order kr and constants C_1 and C_2 .*

Proof of Proposition A.2.1.

In the proof, we will use α to represent $\alpha_{\mathcal{P}}$ without any confusion. The ideas of the proof of Proposition A.2.1 follows from Recht et al. (2010), Candès and Plan (2011).

Denote $S_{kr} = \{X \in \mathbb{R}^{p_1 \times p_2} : \text{rank}(X) \leq kr, \|X\|_F = 1\}$. By Lemma 3.1 in Candès and Plan (2011), for any $\varepsilon > 0$, there exists ε -net S'_{kr} such that $|S'_{kr}| \leq (9/\varepsilon)^{(p_1+p_2+1)kr}$.

For given C_1, C_2 such that $C_1 < 1/(3\alpha^4)$, $C_2 > 1$, we set $C'_1 = \frac{C_1+1/(3\alpha^4)}{2}$, $C'_2 = \frac{C_2+1}{2}$. We can choose δ_0 small enough such that

$$\alpha^2 (2\delta_0 + \delta_0^2) \leq \min \left(\frac{1/(3\alpha^4) - C_1}{2}, \frac{C_2 - 1}{2} \right)$$

then by Lemma A.2.1, for any given $A \in \mathbb{R}^{p_1 \times p_2}$, we have $C'_1 \leq \|\mathcal{X}(A)\|_1/n \leq C'_2$

with probability at least $1 - 2 \exp(-\delta_0^2 n)$. Hence,

$$P(C'_1 \leq \|\mathcal{X}(A)\|_1/n \leq C'_2, \text{ for all } A \in S'_r) \geq 1 - 2(9/\varepsilon)^{kr(p_1+p_2+1)} \cdot \exp(-\delta_0^2 n)$$

Next, we'll estimate the bound of $\|\mathcal{X}(A)\|_1/n$ on the whole set S_{kr} provided that $C'_1 \leq \|\mathcal{X}(A)\|_1/n \leq C'_2$ for all $A \in S'_{kr}$. Define

$$\kappa_1 = \inf_{A \in S_{kr}} \|\mathcal{X}(A)\|_1/n \text{ and } \kappa_2 = \sup_{A \in S_{kr}} \|\mathcal{X}(A)\|_1/n.$$

For any $A \in S_{kr}$, there exists $A' \in S'_{kr}$ such that $\|A - A'\|_F \leq \varepsilon$. So

$$\|\mathcal{X}(A)\|_1/n \leq \|\mathcal{X}(A')\|_1/n + \|\mathcal{X}(A - A')\|_1/n \leq C'_2 \|A\| + \kappa_2 \|A - A'\|_F \leq C'_2 + \kappa_2 \varepsilon$$

$$\|\mathcal{X}(A)\|_1/n \geq \|\mathcal{X}(A')\|_1/n - \|\mathcal{X}(A - A')\|_1/n \geq C'_1 \|A\| - \kappa_2 \|A - A'\|_F \geq C'_1 - \kappa_2 \varepsilon$$

which mean

$$\kappa_2 = \sup_{A \in S_{kr}} \|\mathcal{X}(A)\|_F \leq C'_2 + \varepsilon \kappa_2, \quad \kappa_1 = \inf_{A \in S_{kr}} \|\mathcal{X}(A)\|_F \geq C'_1 - \varepsilon \kappa_2$$

namely, $\kappa_2 \leq C'_2/(1 - \varepsilon)$, $\kappa_1 \geq C'_1 - \varepsilon \kappa_2$. We choose

$$\varepsilon \leq \min \left(\frac{C_2 - 1}{2C_2}, \frac{1/(3\alpha^4) - C_1}{2C_2} \right),$$

by some calculations we can see $\kappa_1 \geq C_1$, $\kappa_2 \leq C_2$.

To sum up, we can choose δ_0, ε only depending on C_1, C_2, α , to ensure

$$C_1 \leq \kappa_1 = \inf_{A \in S_{kr}} \|\mathcal{X}(A)\|_1/n \leq \sup_{A \in S_{kr}} \|\mathcal{X}(A)\|_1/n = \kappa_2 \leq C_2$$

with probability at least $1 - 2(9/\varepsilon)^{kr(p_1+p_2+1)} \exp(-\delta_0^2 n)$. The last step is to estimate

the probability above. We choose $D \geq 8k \log(9/\varepsilon)/\delta_0^2$, then for $n \geq Dr(p_1 + p_2)$, we have

$$\delta_0^2 n/2 \geq 4 \log(9/\varepsilon)kr(p_1 + p_2) \geq \log 2 + 2 \log(9/\varepsilon)kr(p_1 + p_2 + 1),$$

$$1 - 2(9/\varepsilon)^{kr(p_1+p_2+1)} e^{-\delta_0^2 n} = 1 - \exp(-\delta_0^2 n + \log 2 + kr(p_1 + p_2 + 1) \log(9/\varepsilon))$$

$$\geq 1 - \exp(-\delta_0^2 n/2).$$

Finally, we finish the proof of the Theorem by choosing $\delta \leq \delta_0^2/2$. \square

A.2.4 Proof of Theorem 2.2.3, Proposition 2.2.1 and Theorem 2.3.1.

In order to prove the result, we introduce the following technical lemma as an extension of Null Space Property (Lemma A.1.2) from exact low-rank into the approximate low-rank setting.

Lemma A.2.5. *Suppose $A_*, A \in \mathbb{R}^{p_1 \times p_2}$, $R = A_* - A$. If $\|A_*\|_* \leq \|A\|_*$, we have*

$$\|R_{-\max(r)}\|_* \leq \|R_{\max(r)}\|_* + 2\|A_{-\max(r)}\|_* \quad (\text{A.64})$$

The following two lemmas described the separate effect of constraint $\mathcal{Z}_1 = \{z : \|z\|_1/n \leq \lambda_1\}$ and $\mathcal{Z}_2 = \{z : \|\mathcal{X}^*(z)\| \leq \lambda_2\}$ on the estimator.

Lemma A.2.6. *Suppose \mathcal{X} satisfies RUB condition of order kr with constants C_1, C_2 such that $C_1 > C_2/\sqrt{k}$. Assume that $A_*, A \in \mathbb{R}^{p_1 \times p_2}$ satisfy $\|A_*\|_* \leq \|A\|_*$, $\|\mathcal{X}(A_* - A)\|_1/n \leq \lambda_1$. Then we have*

$$\|A_* - A\|_F \leq \frac{2}{C_1 - C_2/\sqrt{k}} \lambda_1 + \left(\frac{3}{\sqrt{k}C_1/C_2 - 1} + \frac{1}{\sqrt{k} - 1} \right) \frac{\|A_{-\max(r)}\|_*}{\sqrt{r}}$$

Lemma A.2.7. *Suppose \mathcal{X} satisfies RUB condition of order kr with constants C_1, C_2 such that $C_1 > C_2/\sqrt{k}$. Assume that \hat{A}^{DS} satisfies $\|\mathcal{X}^* \mathcal{X}(A_* - A)\| \leq \lambda_2$. Then we*

have

$$\|A_* - A\|_F \leq \frac{4}{(C_1 - C_2/\sqrt{k})^2} \cdot \frac{\sqrt{r}\lambda_2}{n} + \left(\frac{5}{\sqrt{k}C_1/C_2 - 1} + \frac{1}{\sqrt{k-1}} + 1 \right) \frac{\|A_{-\max(r)}\|_*}{\sqrt{r}}$$

The proof of Lemma A.2.5, A.2.6 and A.2.7 are listed in the Supplement. Now we prove Theorem 2.2.3 and Proposition 2.2.1. We only need to prove Proposition 2.2.1 since Theorem 2.2.3 is a special case of Proposition 2.2.1. By Lemma A.2.3 and Lemma A.2.4, we have

$$\begin{aligned} P_z(\|z\|_1 \leq \sigma n) &\leq \frac{9}{n}, \\ P_{\mathcal{X},z} \left(\|\mathcal{X}^*(z)\| \geq \sigma \left(12(p_1 + p_2)\sqrt{\log n} + 6\sqrt{2(p_1 + p_2)n} \right) \right) \\ &\leq P_{\mathcal{X}} \left(\|\mathcal{X}^*(z)\| \geq (6(p_1 + p_2)\|z\|_\infty + 6\sqrt{p_1 + p_2}\|z\|_2) \right) \\ &\quad + P_z \left(\|z\|_\infty \geq 2\sigma\sqrt{\log n} \right) + P_z \left(\|z\|_2 \geq \sigma\sqrt{2n} \right) \\ &\leq 2\exp(-(2 - \log 7)(p_1 + p_2)) + \frac{1}{n\sqrt{2\pi \log n}} + \frac{1}{n} \end{aligned}$$

Here $P_{\mathcal{X}}$ (P_z or $P_{\mathcal{X},z}$) means the probability with respect to \mathcal{X} (z or (\mathcal{X}, z)). Hence, we have

$$P(z \in \mathcal{Z}_1 \cap \mathcal{Z}_2) \geq 1 - 2\exp(-(2 - \log 7)(p_1 + p_2)) - \frac{11}{n}.$$

Under the event that $z \in \mathcal{Z}_1 \cap \mathcal{Z}_2$, A is in the feasible set of the programming (2.12), which implies $\|\hat{A}\|_* \leq \|A\|_*$ by the definition of \hat{A} . Moreover, we have

$$\begin{aligned} \|\mathcal{X}(\hat{A} - A)\|_1/n &\leq \|y - \mathcal{X}(A)\|_1/n + \|y - \mathcal{X}(\hat{A})\|_1/n \\ &\leq \|z\|_1/n + \|y - \mathcal{X}(\hat{A})\|_1/n \leq 2\sigma \end{aligned}$$

$$\begin{aligned} \|\mathcal{X}^* \mathcal{X}(\hat{A} - A)\| &\leq \|\mathcal{X}^*(y - \mathcal{X}(\hat{A}))\| + \|\mathcal{X}^*(y - \mathcal{X}(A))\| \\ &\leq \|\mathcal{X}^*(y - \mathcal{X}(\hat{A}))\| + \|\mathcal{X}^*(z)\| \leq 2\eta \end{aligned}$$

On the other hand, suppose $k = 10$, by Theorem 2.2.2, we can have find a uniform

constant C and δ such that if $n \geq Crk(p_1 + p_2)$, \mathcal{X} satisfies RUB of order $10r$ and constants $C_1 = 0.32, C_2 = 1.02$ with probability at least $1 - e^{-n\delta'}$. Hence, we have $D(= Ck)$ and δ' such that if $n \geq Dr(p_1 + p_2)$, \mathcal{X} satisfies RUB of order $10r$ and constants C_1, C_2 satisfying $C_2/C_1 < \sqrt{10}$ with probability at least $1 - e^{-n\delta'}$.

Now under the event that

1. \mathcal{X} satisfies RUB of order $10r$ and constants C_1, C_2 satisfying $C_2/C_1 < \sqrt{10}$,
2. $z \in \mathcal{Z}_1 \cap \mathcal{Z}_2$,

Apply Lemma A.2.6 and Lemma A.2.7 with $A_* = \hat{A}$, we can get (2.17). The probability that these two events both happen is at least $1 - 2 \exp(-(2 - \log 7)(p_1 + p_2)) - \frac{11}{n} - \exp(-\delta'n)$. Set $\delta = \min(2 - \log 7, \delta')$, we finished the proof of Proposition 2.2.1.

For Theorem 2.3.1, the proof is similar. We apply the latter part of Lemma A.2.3 and Lemma A.2.4 and get

$$\begin{aligned}
& P(z \notin \mathcal{Z}_1 \cap \mathcal{Z}_2) \\
& \leq P(\|z\|_1/n > 6\tau) + P\left(\|\mathcal{X}(z)\| > \tau\alpha_{\mathcal{P}}^2 \left(6\sqrt{6n(p_1 + p_2)} + 12\sqrt{\log n(p_1 + p_2)}\right)\right) \\
& \leq P(\|z\|/n > 6\tau) + P(\|z\|_2 > \sqrt{6n\tau}) + P(\|z\|_{\infty} > 2\sqrt{\log n\tau}) \\
& \quad + P_{\mathcal{X}}(\|\mathcal{X}(z)\| > \alpha_{\mathcal{P}}^2(6(p_1 + p_2)\|z\|_{\infty} + 6\sqrt{p_1 + p_2}\|z\|_2)) \\
& \leq \exp\left(-n(6 - 2\sqrt{2\pi})^2/2\right) + \exp(-n/12) + \frac{2}{n} + 2 \exp(-(2 - \log 7)(p_1 + p_2))
\end{aligned}$$

Besides, we choose $k > (3\alpha_{\mathcal{P}}^4)^2$, then we can find $C_1 < 1/(3\alpha_{\mathcal{P}}^4)$ and $C_2 > 1$ such that $C_2/C_1 < \sqrt{k}$. Apply Proposition A.2.1, there exists C, δ' only depending on \mathcal{P}, C_1, C_2 such that if $n \geq Ckr(p_1 + p_2)$, \mathcal{X} satisfies RUB of order kr with constants C_1 and C_2 with probability at least $1 - \exp(-\delta'(p_1 + p_2))$. Note that C_1, C_2 only depends on \mathcal{P} , we can conclude that there exist constants $D(= Ck), \delta'$ only depending on \mathcal{P} such

that if $n \geq Dr(p_1 + p_2)$, \mathcal{X} satisfies RUB of order kr with constants C_1, C_2 satisfying $C_2/C_1 \leq \sqrt{k}$.

Similarly to the proof of Proposition 2.2.1, under the event that

1. \mathcal{X} satisfies RUB of order kr and constants C_1, C_2 satisfying $C_2/C_1 < \sqrt{k}$;
2. $z \in \mathcal{Z}_1 \cap \mathcal{Z}_2$;

we can get (2.28) (we shall note that W_1 depends on \mathcal{P} , so its value can also depend on $\alpha_{\mathcal{P}}$). The probability that those events happen is at least $1 - 2/n - 5 \exp(-\delta(p_1 + p_2))$ for $\delta \leq \min((6 - 2\sqrt{2\pi})^2/2, 1/12, 2 - \log 7, \delta')$. \square

A.2.5 Proof of Theorem 3.3.3.

Without loss of generality, we assume that $p_1 \leq p_2$. We consider the class of rank- r matrices

$$\mathcal{F}_c = \{A \in \mathbb{R}^{p_1 \times p_2} : A_{ij} = 0, \text{ whenever } i \geq r + 1\}$$

namely the matrices with all non-zeros entries in the first r rows. The model (2.1) become

$$y_i = \beta_{1:r}^{(i)T} A_r \gamma^{(i)} + z_i, \quad i = 1, \dots, n$$

where $\beta_{1:r}^{(i)}$ is the vector of the first to the r -th entries of $\beta^{(i)}$. Note that this is a linear regression model with variable $A_r \in \mathbb{R}^{r \times p_2}$, by Lemma 3.11 in Candès and Plan (2011), we have

$$\inf_{\hat{A}} \sup_{A \in \mathcal{F}_c} E \|\hat{A}(y) - A\|_F^2 = \sigma^2 \text{trace} [(\mathcal{X}_r^* \mathcal{X}_r)^{-1}] \quad (\text{A.65})$$

$$\inf_{\hat{A}} \sup_{A \in \mathcal{F}_c} E \|\hat{A}(y) - A\|_F^2 = \infty, \quad \text{when } \mathcal{X}_r^* \mathcal{X}_r \text{ is singular,} \quad (\text{A.66})$$

where $\mathcal{X}_r : \mathbb{R}^{r \times p_2} \rightarrow \mathbb{R}^n$ is the \mathcal{X} constrained on \mathcal{F}_c , Then \mathcal{X}_r sends A_r to

$$\left(\beta_{1:r}^{(1)} A_r \gamma^{(1)}, \dots, \beta_{1:r}^{(n)} A_r \gamma^{(n)} \right)^\top.$$

When $n < p_2 r$, \mathcal{X}_r is singular, hence we have (2.16).

When $n \geq p_2 r$, we can see in order to show (2.15), we only need to show $\text{trace}(\mathcal{X}_r^* \mathcal{X}_r) \geq \frac{p_2 r}{2n}$ with probability at least $1 - 26n^{-1}$. Suppose the singular value of \mathcal{X}_r are $\sigma_i(\mathcal{X}_r)$, $i = 1, \dots, rp_2$, then $\text{trace}(\mathcal{X}_r^* \mathcal{X}_r) = \sum_{i=1}^{p_2 r} \sigma_i^2(\mathcal{X}_r)$.

Suppose \mathcal{X} is ROP while $B \in \mathbb{R}^{r \times p_2}$ is i.i.d. standard Gaussian random matrix (both \mathcal{X} and B_r are random). Then by some calculation we can see

$$E_{B, \mathcal{X}_r} \|\mathcal{X}_r(B)\|_2^2 = n E_{B, \beta, \gamma} (\beta_{1:r}^\top B \gamma)^2 = n \sum_{j=1}^r \sum_{k=1}^{p_2} E (\beta_j B_{jk} \gamma_k)^2 = np_2 r$$

Note (0.20) in the proof of Lemma A.2.1 in the Supplement, we know

$$E \left(\beta_{1:r}^{(i)\top} B \gamma^{(i)} \middle| B \right) \leq 9 \|B\|_F^4.$$

Hence,

$$\begin{aligned} & E \|\mathcal{X}_r(B)\|_2^4 \\ &= \sum_{i=1}^n E \left(\beta_{1:r}^{(i)\top} B \gamma^{(i)} \right)^4 + 2 \sum_{1 \leq i < l \leq n} E \sum_{j=1}^n \left(\beta_{1:r}^{(i)\top} B \gamma^{(i)} \right)^2 \cdot E \sum_{j=1}^n \left(\beta_{1:r}^{(l)\top} B \gamma^{(l)} \right)^2 \\ &= n \cdot 9 E \|B\|_F^4 + n(n-1)(p_2 r)^2 = 9n E (\chi^2(p_2 r))^2 + n(n-1)p_2^2 r^2 \\ &= 9n(p_2^2 r^2 + 2p_2 r) + n(n-1)p_2^2 r^2 \\ &= n^2 p_2^2 r^2 + 2np_2 r(4p_2 r + 9) \leq n^2 p_2^2 r^2 + 26np_2^2 r^2 \end{aligned}$$

Besides,

$$E \|\mathcal{X}_r(B_r)\|_2^2 = E \left(E \left(\|\mathcal{X}_r(B_r)\|_2^2 \middle| \mathcal{X}_r \right) \right) = E \left(\sum_{i=1}^{rp_2} \sigma_i^2(\mathcal{X}_r) \right)$$

$$\begin{aligned}
E\|\mathcal{X}_r(B_r)\|_2^4 &= E\left(E\left(\|\mathcal{X}_r(B_r)\|_2^4 \mid \mathcal{X}_r\right)\right) \\
&= E\left(\sum_{i=1}^{rp_2} 3\sigma_i^4(\mathcal{X}_r) + 2 \sum_{1 \leq i < j \leq rp_2} \sigma_i^2(\mathcal{X}_r)\sigma_j^2(\mathcal{X}_r)\right) \geq E\left(\sum_{i=1}^{rp_2} \sigma_i^2(\mathcal{X}_r)^2\right)^2
\end{aligned}$$

Hence,

$$\begin{aligned}
E\left(\sum_{i=1}^{rp_2} \sigma_i^2(\mathcal{X}_r)^2\right) &= np_2r \\
\text{Var}\left(\sum_{i=1}^{rp_2} \sigma_i^2(\mathcal{X}_r)^2\right) &= E\left(\sum_{i=1}^{rp_2} \sigma_i^2(\mathcal{X}_r)^2\right)^2 - \left(E\sum_{i=1}^{rp_2} \sigma_i^2(\mathcal{X}_r)^2\right)^2 \leq 26np_2^2r^2
\end{aligned}$$

Then by Chebyshev's inequality, we have

$$\sum_{i=1}^{rp_2} \sigma_i^2(\mathcal{X}_r) \leq 2np_2r \tag{A.67}$$

with probability at least $1 - \frac{26np_2^2r^2}{(np_2r)^2} = 1 - \frac{26}{n}$. By Cauchy-Schwarz's inequality, we have

$$\text{trace}\left((\mathcal{X}_r^* \mathcal{X}_r)^{-1}\right) = \sum_{i=1}^{rp_2} \sigma_i^{-2}(\mathcal{X}_r) \geq \frac{(p_2r)^2}{\sum_{i=1}^{rp_2} \sigma_i^2(\mathcal{X}_r)}$$

Therefore, we have

$$\text{trace}\left((\mathcal{X}_r^* \mathcal{X}_r)^{-1}\right) \geq \frac{p_2r}{2n}$$

with probability at least $1 - 26/n$, which shows (2.15).

Finally we consider (2.14). Suppose inequality (A.67) holds, then

$$\begin{aligned}
|\{i : \sigma_i^2(X_r) \geq 4n\}| &\leq \frac{p_2r}{2} \\
\Rightarrow \left|\{i : \sigma_i^{-2}(X_r) \leq \frac{1}{4n}\}\right| &\geq \frac{p_2r}{2} \\
\Rightarrow \left|\{i : \sigma_i^{-2}(X_r) \geq \frac{1}{4n}\}\right| &\geq \frac{p_2r}{2}
\end{aligned} \tag{A.68}$$

By Lemma 3.12 in Candès and Plan (2011), we know

$$\begin{aligned}
& \inf_{\hat{A}} \sup_{A \in \mathcal{F}_c} P_z \left(\|\hat{A} - A\|_F^2 \geq \frac{p_2 r \sigma^2}{16n} \right) \\
&= \inf_{\hat{A}} \sup_{A \in \mathcal{F}_c} E_z 1_{\{x \geq \frac{p_2 r \sigma^2}{16n}\}} (\|\hat{A} - A\|_F^2) \\
&= E_z 1_{\{x \geq \frac{p_2 r \sigma^2}{16n}\}} (\|(\mathcal{X}_r^* \mathcal{X}_r)^{-1} \mathcal{X}_r^*(z)\|_F^2) \\
&= P_z \left(\|(\mathcal{X}_r^* \mathcal{X}_r)^{-1} \mathcal{X}_r^*(z)\|_F^2 \geq \frac{p_2 r \sigma^2}{16n} \right)
\end{aligned}$$

where $1_{\{x \geq \frac{p_2 r \sigma^2}{16n}\}}(\cdot)$ is the indicator function. Note that when $z \stackrel{iid}{\sim} N(0, \sigma^2)$,

$$\|(\mathcal{X}_r^* \mathcal{X}_r)^{-1} \mathcal{X}_r^*(z)\|_F^2$$

is identical distributed as $\sum_{i=1}^{rp_2} \frac{y_i^2}{\sigma_i^2(\mathcal{X}_r)}$, where $y_1, \dots, y_{rp_2} \stackrel{iid}{\sim} N(0, \sigma^2)$, hence,

$$\begin{aligned}
& P \left(\|(\mathcal{X}_r^* \mathcal{X}_r)^{-1} \mathcal{X}_r^*(z)\|_F^2 \leq \frac{p_2 r \sigma^2}{16n} \right) \\
&= P \left(\sum_{i=1}^{rp_2} \frac{y_i^2}{\sigma_i^2(\mathcal{X}_r)} \leq \frac{p_2 r \sigma^2}{16n} \right) \\
&\leq P \left(\sum_{i: \sigma_i^{-2}(\mathcal{X}_r) \geq 1/(4n)} y_i^2 \sigma_i^{-2}(\mathcal{X}_r) \leq \frac{p_2 r \sigma^2}{16n} \right) \\
&\leq P \left(\sum_{i: \sigma_i^{-2}(\mathcal{X}_r) \geq 1/(4n)} \frac{y_i^2}{4n} \leq \frac{p_2 r \sigma^2}{16n} \right) \leq P \left(\chi^2(\lceil \frac{rp_2}{2} \rceil) \leq \frac{p_2 r}{4} \right) \\
&\leq \exp \left(-\frac{rp_2}{32} \right).
\end{aligned}$$

The last inequality is due to the tail bound of χ^2 distribution given by Lemma 1 in Laurent and Massart (2000); the second last inequality is due to (A.68). In summary,

when (A.67) holds, we have

$$\inf_{\hat{A}} \sup_{A \in \mathcal{F}_c} P_z \left(\|\hat{A} - A\|_F^2 \geq \frac{p_2 r \sigma^2}{16n} \right) \leq \exp\left(-\frac{rp_2}{32}\right)$$

Finally since $p_2 \geq (p_1 + p_2)/2$, we showed that with probability at least $1 - 26n^{-1}$, \mathcal{X} satisfies (2.14). \square

A.2.6 Proof of Theorem 2.4.1.

We first introduce the following lemma about the upper bound of $\|z\|_1, \|z\|_2, \|z\|_\infty$.

Lemma A.2.8. *Suppose z is defined as (2.33), then for constants $C_1 > \sqrt{2}$, $M_1 > 1$, we have*

$$\begin{aligned} P \left(\|z\|_1/n \leq \frac{C_1}{n} \sum_{i=1}^n \xi_i^2 \right) &\geq 1 - \frac{9C_1^2 + 6}{n(C_1 - \sqrt{2})^2}, \\ P \left(\frac{C_1}{n} \sum_{i=1}^n \xi_i^2 \leq M_1 C_1 \|\Sigma\|_* \right) &\geq 1 - \frac{9}{n(M_1 - 1)^2}; \end{aligned} \tag{A.69}$$

for constants $C_2 > 1$, $M_2 > 9$,

$$\begin{aligned} P \left(\|z\|_2^2/n \leq \frac{C_2^2 \sum_{i=1}^n \xi_i^4}{n} \right) &\geq 1 - \frac{105(105C_2^4 + 60)}{n(3C_2^2 - 2)^2}, \\ P \left(\frac{C_2^2 \sum_{i=1}^n \xi_i^4}{n} \leq M_2 C_2^2 \|\Sigma\|_*^2 \right) &\geq 1 - \frac{105^2}{n(M_1 - 9)^2}; \end{aligned} \tag{A.70}$$

for constants $C_3 > 1$, $M_3 > 1$,

$$\begin{aligned} P \left(\|z\|_\infty \leq C_3 \log n \max_{1 \leq i \leq n} \xi_i^2 \right) &\geq 1 - \frac{2}{\sqrt{2\pi C_3 \log n}}, \\ P \left(C_3 \log n \max_{1 \leq i \leq n} \xi_i^2 \leq 2C_3 M_3 \log^2 n \left(\sqrt{\|\Sigma\|_*} + \sqrt{2M_3 \log n \|\Sigma\|} \right)^2 \right) &\geq 1 - 2n^{-M_3+1}. \end{aligned} \tag{A.71}$$

The proof of Lemma A.2.8 is listed in the Supplement. The rest of the proof is basically the same as Proposition 2.2.3. Suppose $\mathcal{X}_1, \mathcal{X}_2$ and \tilde{z} are given by (0.36), (0.37) and (0.39) in the Supplement, then $\mathcal{X}_1, \mathcal{X}_2$ are ROP. By Lemma A.2.4,

$$\|\mathcal{X}_1^*(\tilde{z})\| \leq 6 \left(2p\|\tilde{z}\|_\infty + \sqrt{2p}\|\tilde{z}\|_2 \right) \quad (\text{A.72})$$

$$\|\mathcal{X}_2^*(\tilde{z})\| \leq 6 \left(2p\|\tilde{z}\|_\infty + \sqrt{2p}\|\tilde{z}\|_2 \right) \quad (\text{A.73})$$

with probability at least $1 - 4 \exp(-2(2 - \log 7)p)$. Hence there exists $\delta > 0$ such that,

$$\begin{aligned} & P(\Sigma_0 \text{ is NOT in the feasible set of (2.34)}) = P\left(\|z\|_1/n > \eta_1, \text{ or } \|\tilde{\mathcal{X}}^*(\tilde{z})\| > \eta_2\right) \\ & \leq P\left(\|z\|_1/n > \frac{c_1}{n} \sum_{i=1}^n \xi_i^2\right) + P\left(\|\tilde{z}\|_\infty > 2c_3 \log n \max_{1 \leq i \leq n} \xi_i^2\right) + P\left(\|\tilde{z}\|_2 > c_2 \sqrt{2 \sum_{i=1}^n \xi_i^4}\right) \\ & \quad + P\left(\|\tilde{\mathcal{X}}^*(z)\| > 24p\|\tilde{z}\|_\infty + 12\sqrt{2p}\|\tilde{z}\|_2\right) \\ & \leq P\left(\|z\|_1/n > \frac{c_1}{n} \sum_{i=1}^n \xi_i^2\right) + P\left(\|z\|_\infty > c_3 \log n \max_{1 \leq i \leq n} \xi_i^2\right) + P\left(\|z\|_2 > c_2 \sqrt{\sum_{i=1}^n \xi_i^4}\right) \\ & \quad + P\left(\|\mathcal{X}_1^*(\tilde{z})\| > 12p\|\tilde{z}\|_\infty + 6\sqrt{2p}\|\tilde{z}\|_2\right) + P\left(\|\mathcal{X}_2^*(z)\| > 12p\|\tilde{z}\|_\infty + 6\sqrt{2p}\|\tilde{z}\|_2\right) \\ & \leq O(1/n) + 4 \exp(-2(2 - \log 7)p) + \frac{2}{\sqrt{2\pi c_3 \log n}}. \end{aligned}$$

Here we used the fact that $\tilde{\mathcal{X}}^* = \mathcal{X}_1^* + \mathcal{X}_2^*$,

$$\|\tilde{z}\|_2 = \sqrt{\sum_{i=1}^{\lfloor n/2 \rfloor} (z_{2i-1} - z_{2i})^2} \leq \sqrt{\sum_{i=1}^{\lfloor n/2 \rfloor} 2(z_{2i-1}^2 + z_{2i}^2)} \leq \sqrt{2}\|z\|_2,$$

$$\|\tilde{z}\|_\infty = \max_i |z_{2i-1} - z_{2i}| \leq 2 \max_i |z_i| \leq 2\|z\|_\infty.$$

Similarly to the proof of Proposition 2.2.3, since \mathcal{X}_1 is ROP, there exists constants D and δ' such that if $n \geq Drp$, \mathcal{X}_1 satisfies RUB of order $10k$ with constants C_1, C_2

satisfying $C_2/C_1 < \sqrt{10}$ with probability at least $1 - e^{-n\delta'}$.

Now under the event that

1. A is feasible in (2.34),
2. \mathcal{X}_1 satisfies RUB of order $10k$ with constants C_1, C_2 satisfying $C_2/C_1 < \sqrt{10}$,
3. The latter part of (A.69), (A.70) and (A.71) hold for some $M_1 > 1, M_2 > 9, M_3 > 2$,

we can prove (2.36) similarly as the proof of Proposition 2.2.3, which we omit the proof here. \square

A.2.7 Proof of Lemma 2.2.1.

Suppose $\mathcal{X} : \mathbb{R}^{p_1 \times p_2} \rightarrow \mathbb{R}^n$ is given by (2.6), we consider rank-1 matrices:

$$A_1 = e_1^{(p_1)} e_1^{(p_2)T}, \quad A_2 = \frac{\beta^{(1)} \gamma^{(1)}}{\|\beta^{(1)}\|_2 \|\gamma^{(1)T}\|_2}$$

Here $e_1^{(p_1)}, e_2^{(p_2)}$ are the p_1 - and p_2 - dimensional vectors with first entry 1 and others 0, respectively. Then we have $\|A_1\|_F = \|A_2\|_F = 1$.

$$E\|\mathcal{X}(A_1)\|_2^2 = E \sum_{i=1}^n (\beta_1^{(i)})^2 (\gamma_1^{(i)})^2 = n$$

$$\text{Var}\|\mathcal{X}(A_1)\|_2^2 = \text{Var} \sum_{i=1}^n (\beta_1^{(i)})^2 (\gamma_1^{(i)})^2 = n \text{Var} (\beta_1^2 \gamma_1^2) = 8n.$$

By Chebyshev's inequality, we have for all $t > 1$,

$$P(\|\mathcal{X}(A_1)\|_2^2 \geq tn) \leq \frac{8}{n(t-1)^2} \tag{A.74}$$

On the other hand,

$$\begin{aligned}\|\mathcal{X}(A_2)\|_2^2 &= \|\beta^{(1)}\|_2^2 \|\gamma^{(1)}\|_2^2 + \sum_{i=2}^n \left(\frac{\beta^{(i)T} \beta^{(1)}}{\|\beta^{(1)}\|_2} \right)^2 \left(\frac{\gamma^{(i)T} \gamma^{(1)}}{\|\gamma^{(1)}\|_2} \right)^2 \\ &\geq \|\beta^{(1)}\|_2^2 \|\gamma^{(1)}\|_2^2 \sim \chi^2(p_1) \cdot \chi^2(p_2)\end{aligned}$$

By Lemma 1 in Laurent and Massart (2000), we know

$$P(\chi^2(p_1) \geq p_1/2) \leq \exp(-p_1/4), \quad P(\chi^2(p_2) \geq p_2/2) \leq \exp(-p_2/4)$$

Hence,

$$P(\|\mathcal{X}(A_2)\|_2^2 \geq p_1 p_2 / 4) \leq \exp(-p_1/4) + \exp(-p_2/4). \quad (\text{A.75})$$

Combining (A.74) and (A.75), we can see

$$C_2/C_1 \geq \frac{\|\mathcal{X}(A_2)\|_2 / (\sqrt{n} \|A_2\|_F)}{\|\mathcal{X}(A_1)\|_2 / (\sqrt{n} \|A_1\|_F)} = \sqrt{\frac{\|\mathcal{X}(A_2)\|_2^2}{\|\mathcal{X}(A_1)\|_2^2}} \geq \sqrt{\frac{p_1 p_2 / 4}{tn}} \quad (\text{A.76})$$

holds with probability at least $1 - e^{-p_1/4} - e^{-p_2/4} - \frac{8}{n(t-1)^2}$. \square

A.2.8 Proof of Lemma 2.3.1

First, the common used definition for sub-Gaussian distribution of random variable X include the following two.

$$\exists c, C > 0, \text{ such that } P(|X| \geq t) \leq C \exp(-ct^2) \quad (\text{A.77})$$

$$\exists c > 0, \text{ such that } Ee^{tX} \leq \exp(c^2 t^2 / 2) \quad (\text{A.78})$$

Suppose $\alpha_{\mathcal{P}}$ is finite, then we have

$$Ee^{tX} = \sum_{k=0}^{\infty} \frac{t^{2k}}{(2k)!} EX^{2k} \leq \sum_{k=0}^{\infty} \frac{t^{2k}}{(2k)!} \alpha^{2k} (2k-1)!! = \sum_{k=0}^{\infty} \frac{(t\alpha)^{2k}}{2^k k!} = \exp((\alpha t)^2/2)$$

namely X is sub-Gaussian. Now suppose X is sub-Gaussian, then

$$\begin{aligned} EX^{2k} &= (2k) \int_0^{\infty} P(|X| > t) t^{2k-1} dt \leq 2kC \int_0^{\infty} t^{2k-1} e^{-ct^2} dt \\ &= \frac{kC}{c^k} \int_0^{\infty} (ct^2)^{k-1} e^{-ct^2} d(ct^2) = \frac{k!C}{c^k} \leq \left(\frac{\max(C, 1)}{c} \right)^k (2k-1)!! \end{aligned}$$

which implies that $\alpha_{\mathcal{P}} \leq \sqrt{\max(C, 1)/c}$ is finite. \square

A.2.9 Proof of Propositions 2.2.2, 2.2.3 and 2.2.4.

We first show Proposition 2.2.2. Denote $\mathcal{X}_1, \mathcal{X}_2 : \mathbb{R}^{p \times p} \rightarrow \mathbb{R}^{\lfloor \frac{n}{2} \rfloor}$ such that

$$[\mathcal{X}_1]_i(B) = \left(\frac{\beta^{(2i-1)} + \beta^{(2i)}}{\sqrt{2}} \right)^T B \left(\frac{\beta^{(2i-1)} - \beta^{(2i)}}{\sqrt{2}} \right), \quad i = 1, \dots, \lfloor \frac{n}{2} \rfloor \quad (\text{A.79})$$

$$[\mathcal{X}_2](B) = \left(\frac{\beta^{(2i-1)} - \beta^{(2i)}}{\sqrt{2}} \right)^T B \left(\frac{\beta^{(2i-1)} + \beta^{(2i)}}{\sqrt{2}} \right), \quad i = 1, \dots, \lfloor \frac{n}{2} \rfloor \quad (\text{A.80})$$

Note that $\frac{1}{\sqrt{2}} (\beta^{(2i-1)} + \beta^{(2i)})$ and $\frac{1}{\sqrt{2}} (\beta^{(2i-1)} - \beta^{(2i)})$ are independent i.i.d. standard normal samples, so both \mathcal{X}_1 and \mathcal{X}_2 are ROP design (see (2.6)). By Corollary 2.2.1, we know there exists uniform constant C such that whenever $\lfloor \frac{n}{2} \rfloor \geq Cr \cdot 2p$, \mathcal{X}_1 satisfies the following property with probability at least $1 - \exp(-n\delta)$,

$$\forall A \in \{A \in \mathbb{R}^{p \times p} : \text{rank}(A) \leq r\}, \quad A = \arg \min_{B \in \mathbb{R}^{p \times p}} \|B\|_* \quad \text{subject to} \quad \mathcal{X}_1(B) = \mathcal{X}_1(A). \quad (\text{A.81})$$

Now we consider the event that (A.81) holds. We note that for any symmetric matrix B ,

$$[\mathcal{X}_1]_i(B) = \frac{1}{2}\beta^{(2i-1)T} B \beta^{(2i-1)} - \frac{1}{2}\beta^{(2i)T} B \beta^{(2i)} = \frac{1}{2}([\mathcal{X}]_{2i-1}(B) - [\mathcal{X}]_{2i}(B))$$

So $\mathcal{X}(B) = \mathcal{X}(A)$ implies $\mathcal{X}_1(B) = \mathcal{X}_1(A)$ for symmetric A and B . Also, since A is feasible in programming (2.19), we have

$$\begin{aligned} \|A\|_* &\geq \min_{B \in \mathbb{S}^p} \|B\|_* \quad \text{subject to} \quad \mathcal{X}(B) = \mathcal{X}(A) \\ &\geq \min_{B \in \mathbb{R}^{p \times p}} \|B\|_* \quad \text{subject to} \quad \mathcal{X}_1(B) = \mathcal{X}_1(A) \\ &= \|A\|_*, \end{aligned}$$

So we can conclude that A can be exactly recovered by (2.19) given (A.81) holds. In summary, for $n \geq 6Crp$, with probability at least $1 - \exp(-n\delta)$, \mathcal{X} satisfies (A.81), then programming (2.19) can recover all $A \in \mathbb{S}^p$ of rank at most r . \square

Next, we consider Proposition 2.2.3. The idea of the proof is similar to Proposition 2.2.1. Define $\tilde{z} \in \mathbb{R}^{\lfloor \frac{n}{2} \rfloor}$ such that

$$\tilde{z}_i = z_{2i-1} - z_{2i}, \quad i = 1, \dots, \lfloor \frac{n}{2} \rfloor. \quad (\text{A.82})$$

Then $\tilde{z} \stackrel{iid}{\sim} N(0, 2)$. We shall also point out two facts, $\tilde{z} = \tilde{y} - \tilde{\mathcal{X}}(A)$ and $\mathcal{X}^* = \mathcal{X}_1^* + \mathcal{X}_2^*$ (defined as (A.79), (A.80)). By Lemma A.2.3, we know

$$\begin{aligned} P(\|z\|_1/n > \sigma) &\leq \frac{9}{n} \\ P(\|\tilde{z}\|_2 > \sigma\sqrt{2n}) &\leq \frac{1}{\lfloor n/2 \rfloor} \end{aligned}$$

$$P(\|\tilde{z}\|_\infty > 2\sigma\sqrt{2\log n}) \leq \frac{1}{\lfloor n/2 \rfloor}$$

Hence,

$$\begin{aligned}
& P(A \text{ is NOT in the feasible set of programming (2.23)}) \\
&= P\left(\|z\|_1/n > \sigma, \text{ or } \|\tilde{\mathcal{X}}^*(\tilde{z})\| > \eta\right) \\
&= P\left(\|z\|_1/n > \sigma, \text{ or } \|\tilde{\mathcal{X}}^*(\tilde{z})\| > 24\sigma\sqrt{pn} + 48\sigma p\sqrt{2\log n}\right) \\
&\leq P(\|z\|_1/n > \sigma) + P_{\mathcal{X}}(\|\tilde{z}\|_2 > \sigma\sqrt{2n}) + P(\|\tilde{z}\|_\infty > 2\sigma\sqrt{2\log n}) \\
&\quad + P_{\mathcal{X}}\left(\|\tilde{\mathcal{X}}^*(\tilde{z})\| > 24p\|\tilde{z}\|_\infty + 12\sqrt{2p}\|\tilde{z}\|_2\right) \\
&\leq \frac{9}{n} + \frac{2}{\lfloor n/2 \rfloor} + P_{\mathcal{X}}\left(\|\mathcal{X}_1(\tilde{z})\| > 12p\|\tilde{z}\|_\infty + 6\sqrt{2p}\|\tilde{z}\|_2\right) \\
&\quad + P_{\mathcal{X}}\left(\|\mathcal{X}_2(\tilde{z})\| > 12p\|\tilde{z}\|_\infty + 6\sqrt{2p}\|\tilde{z}\|_2\right) \\
&\leq \frac{15}{n} + 4\exp(-2p(2 - \log 7))
\end{aligned}$$

When A is in the feasible set of programming (2.23), we have $\|\hat{A}\|_* \leq \|A\|_*$ and

$$\begin{aligned}
\|\tilde{\mathcal{X}}(\hat{A} - A)\|_1 &= \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \left| [\mathcal{X}]_{2i-1}(\hat{A} - A) - [\mathcal{X}]_{2i}(\hat{A} - A) \right| \\
&\leq \|\mathcal{X}(\hat{A} - A)\|_1 \leq \|y - \mathcal{X}(\hat{A})\|_1 + \|\mathcal{X}(A) - y\|_1 \\
&\leq \|y - \mathcal{X}(\hat{A})\|_1 + \|z\|_1 \leq 2n\sigma
\end{aligned} \tag{A.83}$$

$$\begin{aligned}
\|\tilde{\mathcal{X}}^*\tilde{\mathcal{X}}(\hat{A} - A)\| &\leq \|\tilde{\mathcal{X}}^*(\tilde{y} - \tilde{\mathcal{X}}(\hat{A}))\| + \|\tilde{\mathcal{X}}^*(\tilde{\mathcal{X}}(A) - \tilde{y})\| \\
&= \|\tilde{\mathcal{X}}^*(\tilde{y} - \tilde{\mathcal{X}}(\hat{A}))\| + \|\tilde{\mathcal{X}}^*(\tilde{z})\| \leq 2\eta
\end{aligned} \tag{A.84}$$

Similarly as the proof to Proposition 2.2.1, by Theorem 2.2.2, there exists constant D, δ' such that if $n \geq Drp$, \mathcal{X}_1 satisfies RUB of order $10r$ with constants C_1, C_2 such that $C_2/C_1 < \sqrt{10}$ with probability at least $1 - \exp(-n\delta')$. Now we suppose the following two events happen,

1. \mathcal{X}_1 satisfies RUB of order $10r$ and constants C_1, C_2 satisfying $C_2/C_1 < \sqrt{10}$,
2. A is feasible in the programming (2.23).

Since $\tilde{\mathcal{X}}(B) = 2\mathcal{X}_1(B)$ for any symmetric matrix B , by \mathcal{X}_1 satisfies RUB condition, we have $\tilde{\mathcal{X}}$ satisfies RUB for symmetric matrices of order $10r$ and constants $2C_1, 2C_2$ satisfying $(2C_2)/(2C_1) < \sqrt{10}$. We note that the proof of Lemmas A.2.6 and A.2.7 still apply for $\tilde{\mathcal{X}}$ in the symmetric matrices class, so we can get (2.24) based on (A.83) and (A.84) under those two events happen. Finally the probability that these events happen is at least $1 - 15/n - 4\exp(-p\delta) - \exp(-n\delta)$ for $\delta < \min(2(2 - \log 7), \delta')$, which finished the proof of Proposition 2.2.3. \square

Finally we consider Proposition 2.2.4. Denote $p' = \lfloor p/2 \rfloor$, $r' = \lfloor r/2 \rfloor$. By $r, p \geq 2$, we have $r' \geq r/3, p' \geq p/3$. Define a sub-class of the class rank- r symmetric matrices,

$$\mathcal{G} = \left\{ A \in \mathbb{S}^p : A = \begin{bmatrix} p' & p-p' \\ 0 & B \\ B^T & 0 \end{bmatrix} \begin{matrix} p' \\ p-p' \end{matrix}, B \in \mathbb{R}^{p' \times (p-p')}, \text{rank}(B) \leq r' \right\}$$

we can see $\forall A \in \mathcal{G}$,

$$[\mathcal{X}(A)]_i = \beta^{(i)T} A \beta^{(i)} = 2(\beta_1^{(i)}, \dots, \beta_{p'}^{(i)}) B (\beta_{p'+1}^{(i)}, \dots, \beta_p^{(i)})^T,$$

so in \mathcal{G} the SROP model becomes

$$\frac{y_i}{2} = (\beta_1^{(i)}, \dots, \beta_{p'}^{(i)}) B (\beta_{p'+1}^{(i)}, \dots, \beta_p^{(i)})^T + \frac{z_i}{2}, \quad \frac{z_i}{2} \stackrel{iid}{\sim} N(0, \sigma^2/4)$$

which is an ROP model which we already discussed in section 2.2. We omit the rest of the proof as it can be followed by the proof of Theorem 3.3.3. \square

A.2.10 Proof of Proposition 2.3.1.

The proof can follow from the proof of Proposition 2.2.3 and Theorem 2.3.1 once we can prove that in high probability, \mathcal{X}_1 (defined in (A.79)) satisfies RUB condition with C_1, C_2 such that C_2/C_1 is bounded. This can be proved similarly as Proposition A.2.1, where we only need to edit the proof that we use the following Lemma A.2.9 instead of Lemma A.2.1.

Lemma A.2.9. *Suppose $A \in \mathbb{R}^{p \times p}$ is a fixed matrix (not necessarily symmetric) and \mathcal{X}_1 is given by (A.79). $\beta^{(i)}$ is a set of p -dimensional vectors such that $\stackrel{iid}{\sim} \mathcal{P}$, where \mathcal{P} is some symmetric variance 1 sub-Gaussian distribution except Rademacher ± 1 distribution. Then for $\delta > 0$, we have*

$$\begin{aligned} & \left(\frac{\min^{3/2}(\text{Var}(\mathcal{P}^2)/2, 1)}{3(2\alpha_{\mathcal{P}})^4} - 8\alpha_{\mathcal{P}}^2\delta - 4\alpha_{\mathcal{P}}^2\delta^2 \right) \|A\|_F \leq \frac{\|\mathcal{X}_1(A)\|_1}{\lfloor n/2 \rfloor} \\ & \leq \left(\sqrt{3/2}\alpha_{\mathcal{P}}^2 + 8\alpha_{\mathcal{P}}^2\delta + 4\alpha_{\mathcal{P}}^2\delta^2 \right) \|A\|_F. \end{aligned} \tag{A.85}$$

with probability at least $1 - 2 \exp(-\delta^2 \lfloor n/2 \rfloor)$.

The proof of Lemma A.2.9 is in the Appendix right after this paragraph. Note that provided \mathcal{P} is symmetric and with variance 1, $\text{Var}(\mathcal{P}^2) = 0$ if and only \mathcal{P} is Rademacher ± 1 and A is diagonal, in which the lower bound of (A.85) becomes meaningless. So we only exclude Rademacher ± 1 distribution from the result. \square

Proof of Lemma A.2.9.

The proof of Lemma A.2.9 is basically the same to Lemma A.2.1. We only need to redo two parts of the proof, where there are major differences.

1. Part 1. “Step 1. Even moments of $|\frac{1}{2}(\beta^{(1)} + \beta^{(2)})^T A(\beta^{(1)} - \beta^{(2)})|$.”.

First, based on \mathcal{P} is symmetric and with variance 1, we have $E\mathcal{P}^{2k+1} = 0$, $E\mathcal{P}^{2k} \leq \alpha^{2k} E x^{2k} = \alpha^{2k} (2k - 1)!!$. Then we can calculate that

$$E(\beta_i^{(1)} + \beta_i^{(2)})^{2k+1} = E(\beta_i^{(1)} - \beta_i^{(2)})^{2k+1} = 0$$

$$\begin{aligned} & E(\beta_i^{(1)} + \beta_i^{(2)})^{2k} \\ &= \sum_{l=0}^k \binom{2k}{2l} E(\beta_i^{(1)})^{2l} E(\beta_i^{(2)})^{2(k-l)} \leq \alpha^{2k} \sum_{l=0}^k \binom{2k}{2l} (2l - 1)!! (2(k-l) - 1)!! \\ &= \alpha^{2k} \sum_{l=0}^k \frac{(2k - 1)!! 2^k k!}{2^l l! 2^{k-l} (k-l)!} = \alpha^{2k} (2k - 1)!! \sum_{l=0}^k \binom{k}{l} = 2^k \alpha^{2k} (2k - 1)!! \end{aligned}$$

Similarly, $E(\beta_i^{(1)} - \beta_i^{(2)})^{2k} \leq 2^k \alpha^{2k} (2k - 1)!!$. Next, we can similarly consider the expansion of $E\left(\frac{1}{2}(\beta^{(1)} + \beta^{(2)})A(\beta^{(1)} - \beta^{(2)})\right)^{2k}$, where the non-zero terms can be written as

$$\frac{1}{2^{2k}} \prod_{l=1}^{2k} A_{i_l, j_l} \prod_{i=1}^p E(\beta_i^{(1)} + \beta_i^{(2)})^{2s_i} \prod_{j=1}^p E(\beta_j^{(1)} - \beta_j^{(2)})^{2t_j}$$

Here $s_1 + \dots + s_p = t_1 + \dots + t_p = k$. This term can be bounded as

$$\begin{aligned}
& \left| \frac{1}{2^{2k}} \prod_{l=1}^{2k} A_{i_l, j_l} \prod_{i=1}^p E(\beta_i^{(1)} + \beta_i^{(2)})^{2s_i} (\beta_i^{(1)} - \beta_i^{(2)})^{2t_i} \right| \\
& \leq \frac{1}{2^{2k}} \prod_{l=1}^{2k} |A_{i_l, j_l}| \cdot \prod_{i=1}^p \left(\frac{s_i}{s_i + t_i} E(\beta_i^{(1)} + \beta_i^{(2)})^{2s_i + 2t_i} + \frac{t_i}{s_i + t_i} E(\beta_i^{(1)} - \beta_i^{(2)})^{2s_i + 2t_i} \right) \\
& \leq \frac{1}{2^{2k}} \prod_{l=1}^{2k} |A_{i_l, j_l}| \prod_{i=1}^p 2^{s_i + t_i} \alpha^{2(s_i + t_i)} (2(s_i + t_i) - 1)!! \\
& = \alpha^{4k} \prod_{l=1}^{2k} |A_{i_l, j_l}| \prod_{i=1}^p \frac{(2(s_i + t_i))!}{2^{s_i + t_i} (s_i + t_i)!} \leq \alpha^{4k} \prod_{l=1}^{2k} |A_{i_l, j_l}| \prod_{i=1}^p \frac{((2s_i + 2t_i)!!)^2}{2^{s_i + t_i} s_i! t_i!} \\
& \leq \alpha^{4k} \prod_{l=1}^{2k} |A_{i_l, j_l}| \prod_{i=1}^p \frac{2^{2(s_i + t_i)} ((s_i + t_i)!)^2}{2^{s_i + t_i} s_i! t_i!} \leq \alpha^{4k} 2^{4p} \prod_{l=1}^{2k} |A_{i_l, j_l}| \prod_{i=1}^p \frac{(2s_i)!(2t_i)!}{(2s_i)!!(2t_i)!!} \\
& = \alpha^{4k} 2^{4p} \prod_{l=1}^{2k} |A_{i_l, j_l}| \prod_{i=1}^p E x_i^{2s_i} \cdot \prod_{i=1}^p E y_i^{2t_i}
\end{aligned}$$

Here we assume that $x_i, y_i \stackrel{iid}{\sim} N(0, 1)$. The right hand side of the inequality above is exactly the term in the expansion of $(2\alpha)^{4k} E(x^T A_{abs} y)^{2k}$, where A_{abs} is the element-wise absolute value of A . Therefore, we have

$$E \left(\frac{1}{2} (\beta^{(1)} + \beta^{(2)})^T A (\beta^{(1)} - \beta^{(2)}) \right)^{2k} \leq (2\alpha)^{4k} E[x^T A_{abs} y]^{2k}$$

Now we follow the same argument of the rest part of Step 1 in the proof of Lemma A.2.1, we can prove that

$$E \left(\frac{1}{2} (\beta^{(1)} + \beta^{(2)})^T A (\beta^{(1)} - \beta^{(2)}) \right)^{2k} \leq (2\alpha)^{4k} ((2k - 1)!!)^2 \|A\|_F^{2k} \quad (\text{A.86})$$

2. Part 2. The upper and lower bound of $\mu = E \left| \frac{1}{2} (\beta^{(1)} + \beta^{(2)})^T A (\beta^{(1)} - \beta^{(2)}) \right|$.

To follow the argument of the proof of Step 3 in Lemma A.2.1, we need to derive a new bound for $\mu = E \left| \frac{1}{2} (\beta^{(1)} + \beta^{(2)})^T A (\beta^{(1)} - \beta^{(2)}) \right|$. First, we denote

$M = |\frac{1}{2} (\beta^{(1)} + \beta^{(2)})^T A (\beta^{(1)} - \beta^{(2)})|$, then

$$\begin{aligned}
\mu &= EM \leq \sqrt{EM^2} \\
&= \sqrt{E \left(\frac{1}{2} (\beta^{(1)} + \beta^{(2)})^T A (\beta^{(1)} - \beta^{(2)}) \right)^2} \\
&= \sqrt{\frac{1}{4} \sum_{i,j} E(\beta_i^{(1)} + \beta_i^{(2)})^2 A_{ij}^2 (\beta_j^{(1)} - \beta_j^{(2)})^2} \\
&= \sqrt{\sum_{i \neq j} A_{ij}^2 + \frac{1}{2} \text{Var}(\mathcal{P}^2) \sum_i A_{ii}^2} \leq \sqrt{\sum_{i \neq j} A_{ij}^2 + \frac{3}{2} \alpha_{\mathcal{P}}^4 \sum_i A_{ii}^2} \\
&\leq \sqrt{\frac{3}{2} \alpha_{\mathcal{P}}^2} \|A\|_F
\end{aligned}$$

On the other hand,

$$\begin{aligned}
EM^2 &= E \left(\frac{1}{2} (\beta^{(1)} + \beta^{(2)})^T A (\beta^{(1)} - \beta^{(2)}) \right)^2 \\
&= \sum_{i \neq j} A_{ij}^2 + \frac{1}{2} \text{Var}(\mathcal{P}^2) \sum_i A_{ii}^2 \geq \min\left(\frac{1}{2} \text{Var}(\mathcal{P}^2), 1\right) \|A\|_F^2,
\end{aligned}$$

By (A.86), we also have $EM^4 \leq 9(2\alpha)^8 \|A\|_F^{2k}$. By Hölder's inequality, $EM^2 \leq (EM)^{2/3} (EM^4)^{1/3}$. Hence,

$$\mu \geq \sqrt{\frac{(EM^2)^3}{EM^4}} \geq \frac{\min^{3/2}(\text{Var}(\mathcal{P}^2)/2, 1) \|A\|_F}{3(2\alpha)^4}$$

To sum up, instead of Lemma A.2.2, we have the bound of μ as follows,

$$\frac{\min^{3/2}(\text{Var}(\mathcal{P}^2)/2, 1) \|A\|_F}{3(2\alpha_{\mathcal{P}})^4} \leq \mu \leq \sqrt{3/2} \alpha_{\mathcal{P}}^2 \|A\|_F \quad (\text{A.87})$$

The rest of the proof follows the proof of Lemma A.2.1 with modifications of constants, which we do not go into details. \square

A.2.11 Proofs of Technical Tools

We collect the proofs of technical tools used in the theoretical analysis of Chapter 2 in this section.

Proof of Lemma A.2.1

Without confusion, we simply use α to represent $\alpha_{\mathcal{P}}$ in the proof. Note that we can multiply A by a scale without loss of generality. So we assume throughout the proof that $\|A\|_F = 1$. We'll prove this lemma by steps. First, we show an inequality on the even moments of $|\beta^T A \gamma|$; next, we give a bound on the moment generation function of $|\beta^T A \gamma|$. Finally, we give the desired tail bound.

1. Step 1: Even moments of $|\beta^T A \gamma|$.

Assume that $x = (x_1, \dots, x_{p_1}), y = (y_1, \dots, y_{p_2})$ are two random i.i.d. standard normal distributed vectors. Based on the definition of $\alpha_{\mathcal{P}}$ in (2.26), we know

$$\begin{aligned} E\beta_i^{2k} &= E\gamma_j^{2k} \leq \alpha^{2k} E x_i^{2k} = \alpha^{2k} E y_j^{2k}; \\ E\beta_i^{2k-1} &= E\gamma_j^{2k-1} = E x_i^{2k-1} = E y_j^{2k-1} = 0 \end{aligned} \tag{A.88}$$

Consider the expansion of $E(\beta^T A \gamma)^{2k}$, where the non-zero terms can be written as

$$\prod_{l=1}^{2k} A_{i_l, j_l} \cdot \prod_{i=1}^{p_1} E\beta_i^{2s_i} \cdot \prod_{j=1}^{p_2} E\gamma_j^{2t_j}.$$

Here $s_1 + \dots + s_{p_1} = t_1 + \dots + t_{p_2} = k$. By (A.88), this term can be bounded as

$$\left| \prod_{l=1}^{2k} A_{i_l, j_l} \cdot \prod_{i=1}^{p_1} E\beta_i^{2s_i} \cdot \prod_{j=1}^{p_2} E\gamma_j^{2t_j} \right| \leq \prod_{l=1}^{2k} |A_{i_l, j_l}| \cdot \alpha^{4k} \cdot \prod_{i=1}^{p_1} E x_i^{2s_i} \cdot \prod_{j=1}^{p_2} E y_j^{2t_j}$$

The right hand side is exact the term in the expansion of $\alpha^{4k} E(x^T A_{abs} y)^{2k}$,

where A_{abs} is the the element-wise absolute value of A . Therefore, we have

$$E[\beta^T A \gamma]^{2k} \leq \alpha^{4k} E[x^T A_{abs} y]^{2k}. \quad (\text{A.89})$$

Now we suppose A_{abs} has singular value decomposition

$$A_{abs} = \sum_{i=1}^p a_i u_i v_i^T = U \text{diag}(a) V^T$$

where U, V are orthogonal and $a = (a_1, \dots, a_p)$ is the singular value vector of A_{abs} . A well-known fact is that $\sum_i a_i^2 = \|A_{abs}\|_F^2 = \|A\|_F^2$. Since x, y are standard normal distributed, we can see that $x^T A_{abs} y$ and $x^T \text{diag}(a) y$ has the same distribution. So

$$E[x^T A_{abs} y]^{2k} = E\left[\sum_{i=1}^p a_i x_i y_i\right]^{2k}$$

Next, we note

$$z = \sum_{i=1}^p x_i \frac{a_i y_i}{\sqrt{\sum_{j=1}^p a_j^2 y_j^2}},$$

then z is standard normal distributed and independent of $\sqrt{\sum_{j=1}^p a_j^2 y_j^2}$ since

$$\sum_{i=1}^p \left(\frac{a_i y_i}{\sqrt{\sum_{j=1}^p a_j^2 y_j^2}} \right)^2 = 1.$$

and z given y_1, \dots, y_p is always standard normal distributed.

For integer $k \geq 1$,

$$\begin{aligned}
E[x^T A_{abs} y]^{2k} &= E \left| z \cdot \sqrt{\sum_{j=1}^p a_j^2 y_j^2} \right|^{2k} = E|z|^{2k} \cdot E \left(\sum_{j=1}^p a_j^2 y_j^2 \right)^k \\
&= (2k-1)!! \cdot \sum_{\substack{k_1, k_2, \dots, k_p \geq 0, \\ k_1 + \dots + k_p = k}} \frac{k!}{\prod_{j=1}^p k_j!} \prod_{i=1}^p E(a_i^2 y_i^2)^{k_i} \\
&= \sum_{\substack{k_1, k_2, \dots, k_p \geq 0, \\ k_1 + \dots + k_p = k}} \frac{k!}{\prod_{j=1}^p (k_j)!} \prod_{i=1}^p a_i^{2k_i} \cdot \prod_{i=1}^p (2k_i - 1)!! \\
&\leq \sum_{\substack{k_1, k_2, \dots, k_p \geq 0, \\ k_1 + \dots + k_p = k}} \frac{k!}{\prod_{j=1}^p k_j!} \prod_{i=1}^p a_i^{2k_i} \cdot (2(k_1 + k_2 + \dots + k_p) - 1)!! \\
&= ((2k-1)!!)^2 \cdot \sum_{\substack{k_1, k_2, \dots, k_p \geq 0, \\ k_1 + \dots + k_p = k}} \frac{k!}{\prod_{j=1}^p k_j!} \prod_{i=1}^p (a_i^2)^{k_i} \\
&= ((2k-1)!!)^2 \left(\sum_{i=1}^p a_i^2 \right)^k = ((2k-1)!!)^2 \|A\|_F^{2k}
\end{aligned}$$

Together with (A.89), we have

$$E[\beta^T A \gamma]^{2k} \leq \alpha^{4k} ((2k-1)!!)^2 \|A\|_F^{2k} \quad (\text{A.90})$$

2. Log-moment generation function of $|\beta^T A \gamma|$.

By the bound of the even moments of $|\beta^T A \gamma|$, we can also give the estimate of odd moments, for integer $k \geq 1$,

$$\begin{aligned}
0 \leq E |\beta^T A \gamma|^{2k+1} &\leq \sqrt{E[\beta^T A \gamma]^{2k} \cdot E[\beta^T A \gamma]^{2k+2}} \\
&\leq \alpha^{4k+2} (2k-1)!! \cdot (2k+1)!! \leq \alpha^{4k+2} (2k+1)!
\end{aligned}$$

Also,

$$E [\beta^T A \gamma]^{2k} \leq \alpha^{4k} ((2k-1)!!)^2 \leq \alpha^{4k} (2k)!$$

So for all $k \geq 2$, $E|\beta^T A\gamma|^k \leq \alpha^{2k}k!$. Denote $\mu = E|\beta^T A\gamma|$, then for $0 \leq t < \frac{1}{\alpha^2}$,

$$Ee^{t|\beta^T A\gamma|} = 1 + \sum_{k=1}^{\infty} \frac{t^k}{k!} E|\beta^T A\gamma|^k \leq 1 + t\mu + \sum_{k=2}^{\infty} t^k \alpha^{2k} = 1 + t\mu + \frac{t^2 \alpha^4}{1 - t\alpha^2}$$

For $-\frac{1}{\alpha^2} < t < 0$, we have

$$\begin{aligned} Ee^{t|\beta^T A\gamma|} &= 1 + t\mu + \sum_{k=1}^{\infty} t^{2k} \alpha^{4k} - \sum_{k=1}^{\infty} \frac{|t|^{2k+1}}{(2k+1)!} E|\beta^T A\gamma|^{2k+1} \\ &\leq 1 + t\mu + \frac{t^2 \alpha^4}{1 - |t|^2 \alpha^4} \leq 1 + t\mu + \frac{t^2 \alpha^4}{1 - |t| \alpha^2} \end{aligned}$$

Hence, we have for all $-1/\alpha^2 < t < 1/\alpha^2$,

$$Ee^{t|\beta^T A\gamma|} \leq 1 + t\mu + \frac{t^2 \alpha^4}{1 - |t| \alpha^2}.$$

Note that $\log(1+x) \leq 1+x$ for all $-1 < x < \infty$, we have

$$\log E \exp(t(|\beta^T A\gamma| - \mu)) = \log E \exp(t|\beta^T A\gamma|) - t\mu \leq \frac{t^2 \alpha^4}{1 - |t| \alpha^2}. \quad (\text{A.91})$$

for all $-1/\alpha^2 < t < 1/\alpha^2$.

3. The tail bound of $\|\mathcal{X}(A)\|_1/n$.

Finally, we estimate the tail bound of $\|\mathcal{X}(A)\|_1/n$. Note that

$$\|\mathcal{X}(A)\|_1/n = \left(\sum_{j=1}^n |\beta^{(j)T} A\gamma^{(j)}| \right) / n,$$

based on (A.91), the logarithm of moment generating function of $\|\mathcal{X}(A)\|_1/n$ satisfies

$$\log E \exp(t(\|\mathcal{X}(A)\|_1/n - \mu)) = n \log E \exp\left(\frac{t}{n}(|\beta^T A\gamma| - \mu)\right) \leq \frac{t^2 \alpha^4/n}{1 - |t| \alpha^2/n}$$

By the proof of Lemma 1 in Laurent and Massart (2000), we know $\|\mathcal{X}(A)\|_1/n - \mu$ has the tail bound,

$$P(\|\mathcal{X}(A)\|_1/n - \mu \geq \alpha^2(x/n + 2\sqrt{x/n})) \leq \exp(-x)$$

$$P(\|\mathcal{X}(A)\|_1/n - \mu \leq -\alpha^2(x/n + 2\sqrt{x/n})) \leq \exp(-x)$$

Finally we set $\delta = \sqrt{x/n}$, by Lemma A.2.2, $1/(3\alpha^4) \leq \mu \leq 1$, we finish the proof of Lemma. \square

Proof of Lemma A.2.2

Since \mathcal{P} is symmetric and of variance 1, we have $E\beta_i = E\gamma_j = 0$, $E\beta_i^2 = E\gamma_j^2 = 1$, $E\beta_i^4 = E\gamma_j^4 \leq 3\alpha_{\mathcal{P}}^4$ for all i, j . Then by some expansions and calculations,

$$E(\beta^T A \gamma)^2 = E\left(\sum_{i,j} \beta_i A_{ij} \gamma_j\right)^2 = \sum_{i,j} E\beta_i^2 A_{ij}^2 \gamma_j^2 = \sum_{i,j} A_{ij}^2 = \|A\|_F^2$$

By the first part in the proof of Lemma A.2.1, we have

$$E|\beta^T A \gamma|^4 \leq 9\alpha_{\mathcal{P}}^8 \|A\|_F^4$$

By Hölder's inequality,

$$E|\beta^T A \gamma| \leq \sqrt{E|\beta^T A \gamma|^2} = \|A\|_F$$

which gives the right of the original inequality. For the left, note that

$$E|\beta^T A \gamma|^2 \leq (E|\beta^T A \gamma|)^{2/3} \cdot (E|\beta^T A \gamma|^4)^{1/3}$$

So

$$E|\beta^T A\gamma| \geq \sqrt{\frac{(E|\beta^T A\gamma|^2)^3}{E|\beta^T A\gamma|^4}} \geq \frac{\|A\|_F}{3\alpha_p^4}. \quad \square$$

Proof of Lemma A.2.3

- We first prove the sub-Gaussian part of the lemma. The moment generating function of $|z_i|$ ($t \geq 0$) and z_i^2 ($0 \leq t < 1/(2\gamma)$) satisfy

$$\begin{aligned} Ee^{t|z_i|} &= - \int_0^\infty \exp(t\lambda) dP(|X| \geq \lambda) = 1 + \int_0^\infty P(|X| \geq \lambda) d \exp(t\lambda) \\ &\leq 1 + \int_0^\infty 2t \exp(t\lambda - \lambda^2/(2\gamma^2)) d\lambda \\ &\leq 1 + 2t \exp(t^2\gamma^2/2) \int_0^\infty \exp(-(\lambda - \gamma^2)^2/(2\gamma^2)) d\lambda \\ &\leq 1 + 2t \exp(t^2\gamma^2/2) \sqrt{2\pi}\gamma \\ &\leq \exp(t^2\gamma^2/2) \left(1 + 2t\sqrt{2\pi}\gamma\right) \\ &\leq \exp\left(t^2\gamma^2/2 + 2\sqrt{2\pi}\gamma t\right) \end{aligned}$$

$$\begin{aligned} Ee^{tz_i^2} &= - \int_0^\infty \exp(t\lambda^2) dP(|X| \geq \lambda) = 1 + \int_0^\infty P(|X| \geq \lambda) d \exp(t\lambda^2) \\ &\leq 1 + \int_0^\infty 4t\lambda \exp(-\lambda^2(1/(2\gamma^2) - t)) d\lambda \\ &= 1 + \frac{2t}{1/(2\gamma^2) - t} \end{aligned}$$

Then the moment generating function of $\|z\|_1/n$ and $\|z\|_2^2/n$ satisfies

$$Ee^{t\|z\|_1/n} = (Ee^{t|z_i|/n})^n \leq \exp\left(t^2\gamma^2/(2n) + 2\sqrt{2\pi}t\gamma\right)$$

$$Ee^{t\|z\|_2^2/n} = \left(1 + \frac{2t/n}{1/(2\gamma^2) - t/n}\right)^n \leq \exp\left(\frac{2t}{1/(2\gamma^2) - t/n}\right)$$

Hence for $C \geq 0$,

$$\begin{aligned} P(\|z\|_1/n \geq C) &\leq \frac{E \exp(t\|z\|_1/n)}{\exp(tC)} \leq \exp\left(t^2\gamma^2/(2n) + t(2\sqrt{2\pi}\gamma - C)\right) \\ &= \exp\left(\frac{\gamma^2}{2n} \left(t + \frac{n(2\sqrt{2\pi}\gamma - C)}{\gamma^2}\right)^2 - \frac{n(2\sqrt{2\pi}\gamma - C)^2}{2\gamma^2}\right) \end{aligned}$$

For $C > 2\sqrt{2\pi}\gamma$, we can set $t = \frac{\gamma^2}{n(C-2\sqrt{2\pi}\gamma)}$, then

$$P(\|z\|_1/n \geq C) \leq \exp\left(-\frac{n(C-2\sqrt{2\pi}\gamma)^2}{2\gamma^2}\right)$$

Now we consider the tail bound of $\|z\|_2^2/n$. For $C > 4\gamma^2$,

$$P(\|z\|_2^2/n \geq C) = \frac{E \exp(t\|z\|_2^2/n)}{\exp(tC)} = \exp\left(\frac{2t^2\gamma^2C/n - t(C - 4\gamma^2)}{1 - 2\gamma^2t/n}\right).$$

We set $t = \frac{C-4\gamma^2}{4C\gamma^2/n}$,

$$P(\|z\|_2^2/n \geq C) \leq \exp\left(-\frac{n(C-4\gamma^2)^2}{8\gamma^2C(1-2\gamma^2t/n)}\right) \leq \exp\left(-\frac{n(C-4\gamma^2)^2}{8\gamma^2C}\right)$$

Finally we consider $\|z\|_\infty$,

$$P(\|z\|_\infty \leq C\sqrt{\log n}\gamma) \leq 2n \exp(-C^2 \log n \gamma^2 / (2\gamma^2)) = 2n^{-(C^2/2-1)}$$

- Next, we consider the Gaussian part of the lemma. The bound of $\|z\|_2$ is already given by Lemma 5.1 in Cai et al. (2009). For $\|z\|_1$, we can see $E|z_i|^2 = \sigma^2$,

$$E|z_i| = \frac{\sigma}{\sqrt{2\pi}} \int_0^\infty x e^{-x^2/(2)} dx = \sigma\sqrt{2/\pi}$$

Hence, $E(\|z\|_1/n) = \sigma\sqrt{2/\pi}$, $\text{Var}(\|z\|_1/n) = \text{Var}(|z_i|)/n = (1 - 2/\pi)\sigma^2/n$. By

Chebyshev' inequality,

$$\begin{aligned} P(\|z\|_1/n \geq \sigma) &\leq P\left(\left|\|z\|_1/n - \sigma\sqrt{2/\pi}\right| \geq \sigma(1 - \sqrt{2/\pi})\right) \\ &\leq \frac{\text{Var}(\|z\|_1/n)}{\sigma^2(1 - \sqrt{2/\pi})^2} = \frac{1 + \sqrt{2/\pi}}{(1 - \sqrt{2/\pi})n} \end{aligned}$$

For the bound of $\|z\|_\infty$, we have

$$\begin{aligned} P(\|z\|_\infty \geq 2\sqrt{\log n}\sigma) &\leq \sum_{i=1}^n P(|z_i| \geq 2\sqrt{\log n}\sigma) \\ &\leq n \cdot \frac{2}{2\sqrt{2\pi \log n}} \exp\left(-\frac{1}{2} \cdot 4 \log n\right) = \frac{1}{n\sqrt{2\pi \log n}} \quad \square \end{aligned}$$

Proof of Lemma A.2.4

Again without confusion, we simply use α to represent $\alpha_{\mathcal{P}}$ in the proof. The proof also requires some knowledge of moment generation function and ε -net method. We'll prove by steps.

- **Moment Generation Function of $a^T \mathcal{X}^*(z)b$.** Suppose $a \in \mathbb{R}^{p_1}$, $b \in \mathbb{R}^{p_2}$ are fixed unit vectors. In order to handle the operator norm of $\mathcal{X}^*(z)$, we first consider $a^T \mathcal{X}^*(z)b$. Note that

$$a^T \mathcal{X}^*(z)b = \sum_{i=1}^n z_i a^T \beta^{(i)} \gamma^{(i)T} b \quad (\text{A.92})$$

Denote $X_i = a^T \beta^{(i)}$, $Y_i = b^T \gamma^{(i)}$, then $\{X_i\}_{i=1}^{p_1}$, $\{Y_i\}_{i=1}^{p_2}$ are two independent sets of i.i.d. sub-Gaussian samples. Moreover by $\beta^{(i)}$, $\gamma^{(i)}$ are i.i.d. from symmetric distribution \mathcal{P} , one can show

$$E(X_i)^{2k-1} = E\left(\sum_j a_j \beta_j\right)^{2k-1} = 0,$$

$$\begin{aligned}
E(X_i)^{2k} &= E\left(\sum_{j=1}^{p_1} a_j \beta_j\right)^{2k} = \sum_{k_1+\dots+k_{p_1}=k} \frac{k!}{k_1! \dots k_{p_1}!} \left(\prod_{i=1}^{p_1} a_i^{2k_i} E(\beta_i^{2k_i})\right) \\
&\leq \sum_{k_1+\dots+k_{p_1}=k} \frac{k!}{k_1! \dots k_{p_1}!} \left(\prod_{i=1}^{p_1} a_i^{2k_i} E(x_i^{2k_i}) \alpha^{2k}\right) \\
&= \alpha^{2k} E\left(\sum_{i=1}^{p_1} a_i x_i\right)^{2k} \leq \alpha^{2k} (2k-1)!!
\end{aligned}$$

Here $x_i \stackrel{iid}{\sim} N(0, 1)$. Similarly, $E(Y_i)^{2k-1} = 0$, $E(Y_i)^{2k} \leq \alpha^{2k} (2k-1)!!$. Then for $|t| < 1/\alpha^2$,

$$\begin{aligned}
E \exp(tX_i Y_i) &= \sum_{k=0}^{\infty} \frac{t^k E(X_i Y_i)^k}{k!} \leq \sum_{k=0}^{\infty} \frac{t^{2k} (\alpha^{2k} (2k-1)!!)^2}{(2k)!} \\
&= \sum_{k=0}^{\infty} \frac{(t\alpha^2)^{2k} (2k-1)!!}{2^k k!} = \sum_{k=0}^{\infty} (t\alpha^2)^{2k} \cdot (-1)^k \binom{-1/2}{k} \quad (\text{A.93}) \\
&= \frac{1}{\sqrt{1-t^2\alpha^4}}
\end{aligned}$$

Now for fixed $z \in \mathbb{R}^n$, the logarithm of the moment generating function of $a^T \mathcal{X}^*(z) b$ satisfies

$$\begin{aligned}
\log E \exp(ta^T \mathcal{X}^*(z) b) &= \sum_{i=1}^n \log E \exp(tz_i X_i Y_i) \leq \sum_{i=1}^n -\frac{1}{2} \log(1 - t^2 z_i^2 \alpha^4) \\
&\leq \sum_{i=1}^n \frac{t^2 z_i^2 \alpha^4}{2(1 - t^2 z_i^2 \alpha^4)} \leq \frac{t^2 \|z\|_2^2 \alpha^4}{2(1 - t^2 \|z\|_{\infty}^2 \alpha^4)} \\
&\leq \frac{t^2 \|z\|_2^2 \alpha^4}{2(1 - |t| \|z\|_{\infty} \alpha^2)}
\end{aligned}$$

for any $|t| < 1/(\|z\|_{\infty} \alpha^2)$. Here we used the fact that

$$-\log(1-x) = \sum_{i=1}^{\infty} \frac{x^i}{i} \leq \sum_{i=1}^{\infty} x^i = \frac{x}{1-x}$$

for $0 \leq x < 1$.

- **Tail Bound of $a^T \mathcal{X}^*(z) b$**

By the proof of Lemma 1 in Laurent and Massart (2000), we know for fixed $z \in \mathbb{R}^n$, $a^T \mathcal{X}^*(z)b$ has tail bound: for $x > 0$ and fixed a, b , we have

$$P\left(a^T \mathcal{X}^*(z)b \geq \alpha^2 \left(\|z\|_\infty x + \|z\|_2 \sqrt{2x}\right)\right) \leq \exp(-x);$$

$$P\left(a^T \mathcal{X}^*(z)b \leq -\alpha^2 \left(\|z\|_\infty x + \|z\|_2 \sqrt{2x}\right)\right) \leq \exp(-x).$$

Set $x = C(p_1 + p_2)$, we have

$$\begin{aligned} P\left(|a^T \mathcal{X}^*(z)b| \geq \alpha^2 \left(\|z\|_\infty C(p_1 + p_2) + \|z\|_2 \sqrt{2C(p_1 + p_2)}\right)\right) \\ \leq 2 \exp(-C(p_1 + p_2)). \end{aligned} \quad (\text{A.94})$$

For convenience, We denote

$$T = \alpha^2 \left(\|z\|_\infty (C(p_1 + p_2)) + \|z\|_2 \sqrt{2C(p_1 + p_2)}\right). \quad (\text{A.95})$$

- **ε -net and the upper bound of $\|\mathcal{X}^*(z)\|$.**

In this step, we still fix z . We use the ε -net method to derive the upper bound of $\|\mathcal{X}^*(z)\|_2$, which is given by

$$\|\mathcal{X}^*(z)\|_2 = \sup_{a \in \mathbb{R}^{p_1}, b \in \mathbb{R}^{p_2}} a^T \mathcal{X}^*(z)b$$

From Lemma 2.5 in Vershynin (2011), we can find an ε -net A in the unit sphere of \mathbb{R}^{p_1} , i.e. for all a in the unit sphere of \mathbb{R}^{p_1} , there exists $a' \in A$ such that $\|a' - a\|_2 \leq \varepsilon$. Besides, $|A| \leq (1 + 2/\varepsilon)^{p_1}$. Similarly, there exists ε -net B of the unit ball of \mathbb{R}^{p_2} such that $|B| \leq (1 + 2/\varepsilon)^{p_2}$.

By (A.94), we have

$$P\left(|a^T \mathcal{X}^*(z)b| \geq T, \exists a \in A, b \in B\right) \leq 2(1 + 2/\varepsilon)^{p_1+p_2} \exp(-C(p_1 + p_2)) \quad (\text{A.96})$$

Now we consider under the event that $|a^T \mathcal{X}^*(z)b| \leq T, \forall a \in A, b \in B$. Suppose $\mu = \|\mathcal{X}^*(z)\|_2 = \max_{\|a\|_2=\|b\|_2=1} a^T \mathcal{X}^*(z)b$, $(a^*, b^*) = \arg \max_{a,b} a^T \mathcal{X}^*(z)b$, then we can find $a' \in A, b' \in B$ such that $\|a' - a^*\| \leq \varepsilon, \|b' - b^*\| \leq \varepsilon$. Then,

$$\begin{aligned} \mu &= |a^{*T} \mathcal{X}^*(z)b^*| \\ &= |a'^T \mathcal{X}^*(z)b'| + |(a' - a^*)^T \mathcal{X}^*(z)b'| + |a^{*T} \mathcal{X}^*(z)(b^* - b')| \\ &\leq T + (\|a' - a^*\|_2 + \|b' - b^*\|_2) \cdot \|\mathcal{X}^*(z)\| \leq T + 2\varepsilon\mu \end{aligned}$$

This means $\mu \leq T/(1 - 2\varepsilon)$. Therefore, when $|a^T \mathcal{X}^*(z)b| \leq T, \forall a \in A, b \in B$, we have $\|\mathcal{X}^*(z)\| \leq T/(1 - 2\varepsilon)$.

- Finally, we set $\varepsilon = 1/3$, under the event that

$$\begin{aligned} |a^T \mathcal{X}^*(z)b| &\leq T \\ &= \alpha^2 \left(\|z\|_\infty (C(p_1 + p_2)) + \|z\|_2 \sqrt{2C(p_1 + p_2)} \right), \forall a \in A, b \in B \end{aligned}$$

we have

$$\|\mathcal{X}^*(z)\| \leq T/(1 - 2\varepsilon) \leq 3\alpha^2 \left(\|z\|_\infty C(p_1 + p_2) + \|z\|_2 \sqrt{2C(p_1 + p_2)} \right)$$

By (A.96), the probability that all the event happen is at least

$$1 - 2 \exp(-(C - \log 7)(p_1 + p_2))$$

This finished the proof of the lemma. □

Proof of Lemma A.2.5.

The idea of the proof is originated from Wang and Li (2013), Oymak and Hassibi (2010). We provide the proof here for the completeness of discussion. Note $p = \min(p_1, p_2)$, suppose for any matrix B , $\sigma_i(B)$ is the i -th largest singular value of B . By Lemma 2 in Oymak and Hassibi (2010), we have

$$\begin{aligned}
& \|A_{\max(r)}\|_* + \|A_{-\max(r)}\|_* \\
&= \|A\|_* \geq \|A_*\|_* = \|A - (-R)\|_* \geq \sum_{i=1}^p |\sigma_i(A) - \sigma_i(-R)| \\
&\geq \sum_{i=1}^r (\sigma_i(A) - \sigma_i(R)) + \sum_{i=r+1}^p (\sigma_i(R) - \sigma_i(A)) \\
&= \|A_{\max(r)}\|_* - \|A_{-\max(r)}\|_* + \|R_{-\max(r)}\|_* - \|R_{\max(r)}\|_*
\end{aligned}$$

which implies (A.64). \square

Proof of Lemma A.2.6.

Suppose $R = A_* - A$, then we have

$$\|\mathcal{X}(R)\|_1/n \leq \lambda_1 \tag{A.97}$$

Since $\|A_*\|_* \leq \|A\|_*$. By Lemma A.2.5, we must have (A.64). Suppose $p = \min(p_1, p_2)$ and R has the singular value decomposition, $R = \sum_{i=1}^p \sigma_i u_i v_i^T = U \text{diag}(\vec{\sigma}) V^T$, then $\vec{\sigma}_{-\max(kr)}$ satisfies

$$\begin{aligned}
& \|\vec{\sigma}_{-\max(kr)}\|_\infty \leq \sigma_{kr}, \\
& \|\vec{\sigma}_{-\max(kr)}\|_1 = \|\vec{\sigma}_{-\max(r)}\|_1 - (\sigma_{r+1} + \dots + \sigma_{kr}) \\
& \leq \|\vec{\sigma}_{\max(r)}\|_1 + 2\|A_{-\max(r)}\|_* - (k-1)r\sigma_{kr}
\end{aligned}$$

Set

$$\theta = \max(\sigma_{kr}, (\|\vec{\sigma}_{\max(r)}\|_1 + 2\|A_{-\max(r)}\|_* - (k-1)r\sigma_{kr})/(kr)),$$

then $\|\vec{\sigma}_{-\max(kr)}\|_\infty \leq \theta$, $\|\vec{\sigma}_{-\max(kr)}\|_1 \leq kr\theta$. Similarly to the proof of Theorem 2.2.1, apply Lemma 1.1.1 in Chapter 1, we can get $b^{(i)} \in \mathbb{R}^n$, $\lambda_i \geq 0$, $i = 1, \dots, N$ such that $\vec{\sigma}_{-\max(kr)} = \sum_{i=1}^N \lambda_i b^{(i)}$ and (A.63). Hence,

$$\|b^{(i)}\|_2 \leq \sqrt{\|b^{(i)}\|_1 \cdot \|b^{(i)}\|_\infty} \leq \sqrt{\theta(\|\vec{\sigma}_{\max(r)}\|_1 + 2\|A_{-\max(r)}\|_* - (k-1)r\sigma_{kr})}$$

If $\theta = \sigma_{kr}$, we can optimize over σ_{kr} in the inequality,

$$\begin{aligned} \|b^{(i)}\|_2 &\leq \sqrt{\sigma_{kr}(\|\vec{\sigma}_{\max(r)}\|_1 + 2\|A_{-\max(r)}\|_* - (k-1)r\sigma_{kr})} \\ &\leq \frac{\|\vec{\sigma}_{\max(r)}\|_1 + 2\|A_{-\max(r)}\|_*}{2\sqrt{r(k-1)}}; \end{aligned}$$

if $\theta = (\|\vec{\sigma}_{\max(r)}\|_1 + 2\|A_{-\max(r)}\|_* - (k-1)r\sigma_{kr})/(kr)$, we have

$$\begin{aligned} \|b^{(i)}\|_2 &\leq \frac{\|\vec{\sigma}_{\max(r)}\|_1 + 2\|A_{-\max(r)}\|_* - (k-1)r\sigma_{kr}}{\sqrt{kr}} \\ &\leq \frac{\|\vec{\sigma}_{\max(r)}\|_1 + 2\|A_{-\max(r)}\|_*}{\sqrt{kr}}. \end{aligned} \tag{A.98}$$

Since $k \geq 2$, we always have (A.98). Next, we define $B_i = U \text{diag}(b^{(i)}) V^T$, then the rank of B_i are at most kr , $\sum_{i=1}^N \lambda_i B_i = R_{-\max(kr)}$ and $\|B_i\|_F = \|b^{(i)}\|_2$. Then

$$\begin{aligned}
\lambda_1 &\geq \|\mathcal{X}(R)\|_1/n \geq \|\mathcal{X}(R_{\max(kr)})\|_1/n - \|\mathcal{X}(R_{-\max(kr)})\|_1/n \\
&\geq C_1 \|R_{\max(kr)}\|_F - \sum_{i=1}^N \|\mathcal{X}(\lambda_i B_i)\|_1/n \\
&\geq C_1 \|R_{\max(kr)}\|_F - \sum_{i=1}^N \lambda_i C_2 \|B_i\|_F \\
&\geq C_1 \|R_{\max(kr)}\|_F - C_2 \frac{\|R_{\max(r)}\|_* + 2\|A_{-\max(r)}\|_*}{\sqrt{kr}} \\
&\geq C_1 \|R_{\max(kr)}\|_F - \frac{C_2}{\sqrt{k}} \|R_{\max(kr)}\|_F - \frac{2C_2 \|A_{-\max(r)}\|_*}{\sqrt{kr}},
\end{aligned} \tag{A.99}$$

where the last inequality is due to $\|R_{\max(kr)}\|_F \geq \|R_{\max(r)}\|_F \geq \sqrt{r} \|R_{\max(r)}\|_*$. Therefore,

$$\|R_{\max(kr)}\|_F \leq \frac{\lambda_1}{C_1 - C_2/\sqrt{k}} + \frac{2\|A_{-\max(r)}\|_*}{\sqrt{r}(\sqrt{k}C_1/C_2 - 1)} \tag{A.100}$$

Finally,

$$\begin{aligned}
\|R_{-\max(kr)}\|_F &= \|\vec{\sigma}_{-\max(kr)}\|_2 \leq \sqrt{\|\vec{\sigma}_{-\max(kr)}\|_1 \cdot \|\vec{\sigma}_{-\max(kr)}\|_\infty} \\
&\leq \sqrt{\sigma_{kr} \cdot (\|\vec{\sigma}_{-\max(r)}\|_1 - r(k-1)\sigma_{kr})} \\
&\leq \frac{\|\vec{\sigma}_{-\max(r)}\|_1}{2\sqrt{r(k-1)}} \leq \frac{\|\vec{\sigma}_{\max(r)}\|_1 + 2\|A_{-\max(r)}\|_*}{2\sqrt{r(k-1)}} \\
&\leq \frac{\|R_{\max(r)}\|_F}{2\sqrt{k-1}} + \frac{\|A_{-\max(r)}\|_*}{\sqrt{r(k-1)}}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\|R\|_F &= \sqrt{\|R_{\max(kr)}\|_F^2 + \|R_{-\max(kr)}\|_F^2} \\
&\leq \sqrt{\|R_{\max(kr)}\|_F^2 + \left(\frac{\|R_{\max(r)}\|_F}{2\sqrt{k-1}} + \frac{\|A_{-\max(r)}\|_*}{\sqrt{r(k-1)}} \right)^2} \\
&\leq \sqrt{1 + \frac{1}{4(k-1)}} \|R_{\max(kr)}\|_F + \frac{\|A_{-\max(r)}\|_*}{\sqrt{r(k-1)}} \\
&\leq \left(1 + \frac{1}{8(k-1)} \right) \|R_{\max(kr)}\|_F + \frac{\|A_{-\max(r)}\|_*}{\sqrt{r(k-1)}} \\
&\leq \frac{2}{C_1 - C_2/\sqrt{k}} \lambda_1 + \left(\frac{3}{\sqrt{k}C_1/C_2 - 1} + \frac{1}{\sqrt{k-1}} \right) \frac{\|A_{-\max(r)}\|_*}{\sqrt{r}}
\end{aligned} \tag{A.101}$$

Proof of Lemma A.2.7.

The proof of this theorem is similar to the proof of Lemma A.2.6. Suppose $R = A_* - A$.

In this case we have

$$\|\mathcal{X}^* \mathcal{X}(R)\| \leq \lambda_2 \tag{A.102}$$

instead of (A.97). Besides, since $\|A_*\|_* \leq \|A\|_*$ and Lemma A.2.5, we still have

(A.64). With the similar argument as (A.99), we have

$$\begin{aligned}
\lambda_2 \|R\|_* &\geq \langle R, \mathcal{X}^* \mathcal{X}(R) \rangle = \|\mathcal{X}(R)\|_2^2 \geq \|\mathcal{X}(R)\|_1^2 / n \\
&\geq n \left(C_1 \|R_{\max(kr)}\|_F - \frac{C_2}{\sqrt{k}} \|R_{\max(kr)}\|_F - \frac{2C_2 \|A_{-\max(r)}\|_*}{\sqrt{kr}} \right)_+^2
\end{aligned} \tag{A.103}$$

Here $(x)_+$ means $\max(x, 0)$. Besides,

$$\begin{aligned}
\lambda_2 \|R\|_* &\leq \lambda_2 (\|R_{\max(r)}\|_* + (\|R_{\max(r)}\|_* + 2\|A_{-\max(r)}\|_*)) \\
&\leq 2\lambda_2 (\sqrt{r} \|R_{\max(r)}\|_F + \|A_{-\max(r)}\|_*) \\
&\leq 2\lambda_2 (\sqrt{r} \|R_{\max(kr)}\|_F + \|A_{-\max(r)}\|_*)
\end{aligned}$$

Suppose $x = \|R_{\max(kr)}\|_F$, $y = \|A_{-\max(r)}\|_*/\sqrt{r}$. Based on the previous two inequalities, we have

$$n \left((C_1 - C_2/\sqrt{k})x - \frac{2C_2}{\sqrt{k}}y \right)_+^2 \leq 2\sqrt{r}(x+y)\lambda_2$$

When $x \geq \frac{2C_2y}{\sqrt{k}(C_1 - C_2/\sqrt{k})}$, the inequality above leads to

$$n(C_1 - C_2/\sqrt{k})^2 x^2 - \left(2n(C_1 - C_2/\sqrt{k}) \frac{2C_2}{\sqrt{k}}y + 2\sqrt{r}\lambda_2 \right) x - 2\sqrt{r}\lambda_2 y \leq 0. \quad (\text{A.104})$$

Note that for second order inequality $ax^2 - bx - c \leq 0$, $a > 0$, $b, c \geq 0$, we have $x \leq \frac{b + \sqrt{b^2 + 4ac}}{2a} \leq b/a + \sqrt{c/a}$. Hence we can get an upper bound of x from (A.104).

$$\begin{aligned} x &\leq \frac{2\sqrt{r}\lambda_2}{n(C_1 - C_2/\sqrt{k})^2} + \frac{4C_2/\sqrt{k}y}{(C_1 - C_2/\sqrt{k})} + \frac{\sqrt{2\sqrt{r}\lambda_2 y}}{\sqrt{n}(C_1 - C_2/\sqrt{k})} \\ &\leq \frac{2\sqrt{r}\lambda_2}{n(C_1 - C_2/\sqrt{k})^2} + \frac{4C_2/\sqrt{k}y}{(C_1 - C_2/\sqrt{k})} + \frac{\sqrt{r}\lambda_2}{n(C_1 - C_2/\sqrt{k})^2} + \frac{1}{2}y. \end{aligned}$$

Hence whenever $x \geq \frac{2C_2y}{\sqrt{k}(C_1 - C_2/\sqrt{k})}$ or not,

$$\begin{aligned} &\|R_{\max(kr)}\|_F = x \\ &\leq \max \left\{ \frac{2C_2y}{\sqrt{k}(C_1 - C_2/\sqrt{k})}, \frac{3\sqrt{r}\lambda_2}{n(C_1 - C_2/\sqrt{k})^2} + \left(\frac{4C_2/\sqrt{k}}{(C_1 - C_2/\sqrt{k})} + \frac{1}{2} \right) y \right\} \quad (\text{A.105}) \\ &\leq \frac{3\sqrt{r}\lambda_2}{n(C_1 - C_2/\sqrt{k})^2} + \left(\frac{4C_2/\sqrt{k}}{(C_1 - C_2/\sqrt{k})} + \frac{1}{2} \right) \frac{\|A_{-\max(r)}\|_*}{\sqrt{r}}. \end{aligned}$$

Finally, similarly to (A.101) in Lemma A.2.6, we can get the upper bound of $\|R\|_F$,

$$\begin{aligned} \|R\|_F &\leq \left(1 + \frac{1}{8k-8} \right) \|R_{\max(kr)}\|_F + \frac{\|A_{-\max(r)}\|_*}{\sqrt{r}(k-1)} \\ &\leq \frac{4}{(C_1 - C_2/\sqrt{k})^2} \cdot \frac{\sqrt{r}(\varepsilon + \eta)}{n} + \left(\frac{5}{\sqrt{k}C_1/C_2 - 1} + \frac{1}{\sqrt{k-1}} + 1 \right) \frac{\|A_{-\max(r)}\|_*}{\sqrt{r}} \end{aligned}$$

which finished the proof of lemma A.2.7. \square

Proof of Lemma A.2.8.

Note that $\xi_i \mid \beta^{(i)} \sim N(0, \beta^{(i)T} \Sigma \beta^{(i)})$, we can assume that

$$\xi_i^2 = \beta^{(i)T} \Sigma \beta^{(i)} \cdot Z_i, \quad (\text{A.106})$$

where $Z_i \stackrel{iid}{\sim} (N(0, 1))^2$ and $Z_i, \beta^{(i)}$ are independent. Based on the definition of z in (2.33), we have

$$z_i = y_i - [\mathcal{X}(\Sigma_0)]_i = \xi_i^2 - \beta^{(i)T} \Sigma \beta^{(i)} = \beta^{(i)T} \Sigma \beta^{(i)} (Z_i - 1), \quad i = 1, \dots, n. \quad (\text{A.107})$$

We also denote

$$Q_1 = \frac{C_1}{n} \sum_{i=1}^n \xi_i^2, \quad Q_2 = \frac{C_2^2}{n} \sum_{i=1}^n \xi_i^4, \quad Q_3 = C_3 \cdot \log n \max_{1 \leq i \leq n} \xi_i^2.$$

- We'll first consider the former part of (A.69), (A.70) and (A.71). Suppose $Z \sim (N(0, 1))^2$. It is well known that the non-central m -th moment of Z is $(2m - 1)!!$, so we have

$$E(C_1 Z - |Z - 1|) \geq C_1 - \sqrt{E|Z - 1|^2} = C_1 - \sqrt{2} \quad (\text{A.108})$$

$$E(C_1 Z - |Z - 1|)^2 \leq E(C_1 Z)^2 + E(Z - 1)^2 = 3C_1^2 + 2 \quad (\text{A.109})$$

$$E(C_2^2 Z^2 - (Z - 1)^2) = 3C_2^2 - 2 \quad (\text{A.110})$$

$$E(C_2^2 Z^2 - (Z - 1)^2)^2 \leq C_2^4 E(Z)^4 + E(Z - 1)^4 = 105C_2^4 + 60 \quad (\text{A.111})$$

Next we consider the random quadratic form of Σ . Suppose $\beta = (\beta_1, \dots, \beta_p) \stackrel{iid}{\sim} N(0, 1)$, $X_1, \dots, X_p \stackrel{iid}{\sim} N(0, 1)$, $\lambda_1(\Sigma), \dots, \lambda_p(\Sigma)$ are the eigenvalues of Σ . Since Σ

is positive definite, we have

$$E\beta^T\Sigma\beta = \text{tr}(\Sigma) = \|\Sigma\|_* \quad (\text{A.112})$$

$$\begin{aligned} E(\beta^T\Sigma\beta)^2 &= E\left(\sum_i \beta_i^2\Sigma_{ii} + 2\sum_{i<j} \Sigma_{ij}\beta_i\beta_j\right)^2 \\ &= \sum_i \Sigma_{ii}^2 E\beta_i^4 + 2\sum_{i<j} \Sigma_{ii}\Sigma_{jj} E\beta_i^2\beta_j^2 + \sum_{i<j} 4\Sigma_{ij}^2 E\beta_i^2\beta_j^2 \\ &= 2\left(\sum_{i,j} \Sigma_{ij}^2\right) + \left(\sum_{ii} \Sigma_i\right)^2 = 2\|\Sigma\|_F^2 + \|\Sigma\|_*^2 \end{aligned}$$

Hence,

$$\|\Sigma\|_*^2 \leq E(\beta^T\Sigma\beta)^2 \leq 3\|\Sigma\|_*^2 \quad (\text{A.113})$$

$$\begin{aligned} E(\beta^T\Sigma\beta)^4 &= E\left(\sum_{i=1}^p \lambda_i(\Sigma)X_i^2\right)^4 \\ &= \sum_{1\leq i,j,s,t\leq p} \lambda_i(\Sigma)\lambda_j(\Sigma)\lambda_s(\Sigma)\lambda_t(\Sigma) E X_i^2 X_j^2 X_s^2 X_t^2 \\ &\leq \sum_{1\leq i,j,s,t\leq p} \lambda_i(\Sigma)\lambda_j(\Sigma)\lambda_s(\Sigma)\lambda_t(\Sigma) 7!! = 105\|\Sigma\|_*^4 \end{aligned} \quad (\text{A.114})$$

Then we consider $C_1\xi_i^2 - |z_i|$ and $C_2^2\xi_i^4 - z_i^2$. By (A.106) and (A.107), we have

$$C_1\xi_i^2 - |z_i| = \beta^{(i)T}\Sigma\beta^{(i)} \cdot (C_1Z_i - |Z_i - 1|),$$

$$C_2^2\xi_i^4 - z_i^2 = (\beta^{(i)T}\Sigma\beta^{(i)})^2 (C_2^2Z^2 - (Z - 1)^2),$$

while $\beta^{(i)}$ and Z_i are independent in the equation above. By (A.108)-(A.114), we obtain an estimation of the first and second moment of these two quantities as

$$E(C_1\xi_i^2 - |z_i|) \geq (C_1 - \sqrt{2})\|\Sigma\|_* \quad (\text{A.115})$$

$$\begin{aligned} \text{Var} (C_1 \xi_i^2 - |z_i|) &\leq E (C_1 \xi_i^2 - |z_i|)^2 \\ &= E (C_1 Z_i - |Z_i - 1|)^2 E (\beta^{(i)T} \Sigma \beta^{(i)})^2 \leq (9C_1^2 + 6) \|\Sigma\|_*^2 \end{aligned} \quad (\text{A.116})$$

$$E (C_2^2 \xi_i^4 - z_i^2) \geq (3C_2^2 - 2) \cdot \|\Sigma\|_*^2 \quad (\text{A.117})$$

$$\text{Var} (C_2^2 \xi_i^4 - z_i^2) \leq E (C_2^2 \xi_i^4 - z_i^2)^2 \leq 105(105C_2^4 + 60) \|\Sigma\|_*^4 \quad (\text{A.118})$$

We note that $Q_1 - \|z\|_1/n$ and $Q_2 - \|z\|_2^2/n$ are the average of n i.i.d. copy of $C_1 \xi_i^2 - |z_i|$ and $C_2^2 \xi_i^4 - z_i^2$. We can immediately get an estimation of the mean and variance of $Q_1 - \|z\|_1/n$ and $Q_2 - \|z\|_2^2/n$ based on (A.115)-(A.118). Finally, by Chebyshev's inequality,

$$P(Q_1 \leq \|z\|_1/n) \leq \frac{\text{Var}(Q_1 - \|z\|_1/n)}{(E(Q_1 - \|z\|_1/n))^2} \leq \frac{9C_1^2 + 6}{n(C_1 - \sqrt{2})^2}$$

$$P\left(Q_2 \leq \frac{\|z\|_2^2}{n}\right) \leq \frac{\text{Var}(Q_2 - \|z\|_2^2/n)}{(E(Q_2 - \|z\|_2^2/n))^2} \leq \frac{105(105C_2^4 + 60)}{n(3C_2^2 - 2)^2}$$

Since $C_3 > 1$ and $n \geq 3$, we know $C_3 \log n Z \geq Z - 1$ with probability 1. Suppose $i_0 = \arg \max_i \beta^{(i)T} \Sigma \beta^{(i)}$, then

$$\begin{aligned} &P(Q_3 \leq \|z\|_\infty) \\ &\leq P\left(\max_i (C_3 \log n (\beta^{(i)T} \Sigma \beta^{(i)}) Z_i) \leq \max_i ((\beta^{(i)T} \Sigma \beta^{(i)})(Z_i - 1))\right) \\ &\quad + P\left(\max_i (C_3 \log n (\beta^{(i)T} \Sigma \beta^{(i)}) Z_i) \leq \max_i ((\beta^{(i)T} \Sigma \beta^{(i)})(1 - Z_i))\right) \\ &\leq 0 + P\left(\max_i (C_3 \log n (\beta^{(i)T} \Sigma \beta^{(i)}) Z_i) \leq \max_i (\beta^{(i)T} \Sigma \beta^{(i)})\right) \\ &\leq P(C_3 \log n \beta^{(i_0)T} \Sigma \beta^{(i_0)} Z_{i_0} \leq \beta^{(i_0)T} \Sigma \beta^{(i_0)}) \\ &\leq P\left(Z_{i_0} \leq \frac{1}{C_3 \log n}\right) = P\left(|N(0, 1)| \leq \frac{1}{\sqrt{C_3 \log n}}\right) \leq \frac{2}{\sqrt{2\pi C_3 \log n}} \end{aligned}$$

- Then we consider the latter part of (A.69)-(A.71). We can do similar calcula-

tions as the first part of the proof and get

$$\begin{aligned}
EQ_1 &= C_1 E \xi_i^2 = C_1 E (\beta^{(i)T} \Sigma \beta^{(i)}) = C_1 \|\Sigma\|_* \\
\text{Var}(Q_1) &= \frac{C_1^2}{n} \text{Var} \xi_i^2 \leq \frac{C_1^2}{n} E \xi_i^4 = \frac{C_1^2}{n} E (\beta^{(i)T} \Sigma \beta^{(i)})^2 E Z^2 \leq \frac{9C_1^2}{n} \|\Sigma\|_*^2 \\
EQ_2 &= C_2^2 E Z^2 \cdot E (\beta^{(i)T} \Sigma \beta^{(i)})^2 \leq 9C_2^2 \|\Sigma\|_*^2 \\
\text{Var}(Q_2) &= \frac{1}{n} \text{Var} (C_2^2 \xi_i^4) \leq \frac{C_2^4}{n} E \xi_i^8 \\
&= \frac{C_2^4}{n} E Z^4 \cdot E (\beta^{(i)T} \Sigma \beta^{(i)})^4 \leq \frac{105^2 C_2^4}{n} \|\Sigma\|_*^4
\end{aligned}$$

So by Chebyshev's inequality,

$$P(Q_1 \geq M_1 C_1 \|\Sigma\|_*) \leq P(Q_1 - EQ_1 \geq (M_1 - 1)C_1 \|\Sigma\|_*) \leq \frac{9}{(M_1 - 1)^2 n}$$

$$\begin{aligned}
P(Q_2 \geq M_2 C_2^2 \|\Sigma\|_*^2) &\leq P(Q_2 - EQ_2 \geq (M_2 - 9)C_2^2 \|\Sigma\|_*^2) \\
&\leq \frac{\text{Var}(Q_2)}{(M_2 - 9)^2 C_2^4 \|\Sigma\|_*^4} \leq \frac{105^2}{n(M_2 - 9)^2}
\end{aligned}$$

which provide the latter part of (A.69) and (A.70). Finally we note that $\xi_i^2 = (\beta^{(i)T} \Sigma \beta^{(i)}) Z_i$. By Lemma 1 in Laurent and Massart (2000) and the fact that $\|\Sigma\|_* = \sum_i \lambda_i(\Sigma)$, $\|\Sigma\| = \max_i \lambda_i(\Sigma)$, $\|\Sigma\|_F = \sqrt{\sum_i \lambda_i^2(\Sigma)}$, $\|\Sigma\|_F \leq$

$\sqrt{\|\Sigma\|_* \cdot \|\Sigma\|}$, we have

$$\begin{aligned}
& P\left(\beta^T \Sigma \beta \geq \left(\sqrt{\|\Sigma\|_*} + \sqrt{2M_3 \log n \|\Sigma\|}\right)^2\right) \\
&= P\left(\sum_{i=1}^n \lambda_i(\Sigma) X_i^2 \geq \|\Sigma\|_* + 2\sqrt{2M_3 \log n \|\Sigma\|_* \|\Sigma\|} + 2M_3 \log n \|\Sigma\|\right) \\
&\leq P\left(\sum_{i=1}^n \lambda_i(\Sigma) (X_i^2 - 1) \geq 2\sqrt{2M_3 \log n \sum_{i=1}^n \lambda_i^2(\Sigma)} + 2M_3 \log n \max_i \lambda_i(\Sigma)\right) \\
&\leq n^{-M_3}
\end{aligned} \tag{A.119}$$

and

$$P(Z \geq 2M_3 \log n) \leq \exp(-2M_3 \log n/2) = n^{-M_3}.$$

Hence,

$$P\left(C_3 \log n \xi_i^2 \geq 2C_3 M_3 \log^2 n \left(\sqrt{\|\Sigma\|_*} + \sqrt{2M_3 \log n \|\Sigma\|}\right)^2\right) \leq 2n^{-M_3},$$

and consequently

$$P\left(C_3 \log n \max_{1 \leq i \leq n} \xi_i^2 \geq 2C_3 M_3 \log^2 n \left(\sqrt{\|\Sigma\|_*} + \sqrt{2M_3 \log n \|\Sigma\|}\right)^2\right) \leq 2n^{-M_3+1},$$

which gives the right side of (A.71). \square

A.3 Supplement for Chapter 3

In this Appendix we provide additional simulation results and the proofs of the main theorems. Some key technical tools used in the proofs of the main results are also developed and proved.

A.3.1 Additional Simulation Results

We consider the effect of the number of the observed rows and columns on the estimation accuracy. We let $p_1 = p_2 = 1000$, let the singular values of A be $\{j^{-1}, j = 1, 2, \dots\}$ and let m_1 and m_2 vary from 10 to 210. The singular spaces U and V are again generated randomly from the Haar measure. The estimation errors of \hat{A}_{22} from Algorithm 2 with row thresholding and $T_R = 2\sqrt{p_1/m_1}$ over different choices of m_1 and m_2 are shown in Figure A.1. As expected, the average loss decreases as m_1 or m_2 grows.

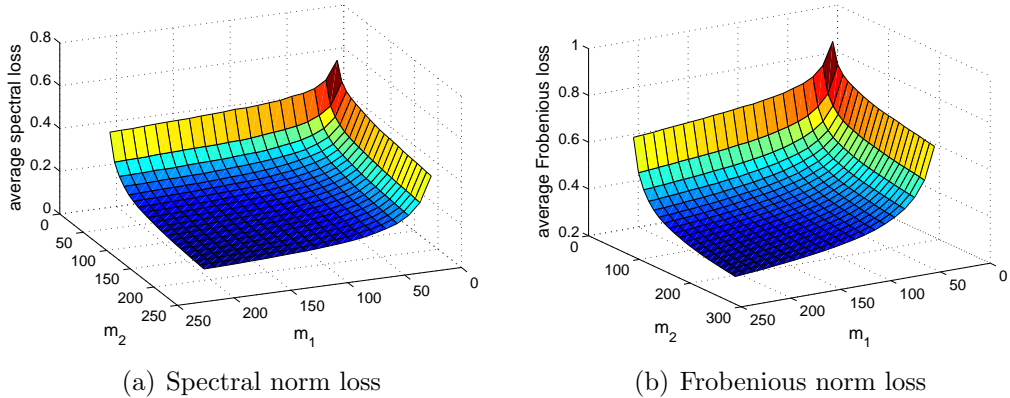


Figure A.1: Losses for the settings with singular values of A being $\{j^{-1}, j = 1, 2, \dots\}$, $p_1 = p_2 = 1000$, $m_1, m_2 = 10, \dots, 210$.

Another interesting fact is that the average loss is approximately symmetric with respect to m_1 and m_2 . This implies that even with different numbers of observed rows and columns, Algorithm 2 has similar performance with row thresholding or column

thresholding.

We are also interested in the performance of Algorithm 2 as p_1 and the ratio m_1/p_1 vary. To this end, we consider the setting where $p_2 = 1000$, $m_2 = 50$, and the singular values of A are chosen as $\{j^{-1}, j = 1, 2, \dots\}$. The results are shown in Figure A.2. It can be seen that when m_1/p_1 increases, the recovery is generally more accurate; when m_1/p_1 is kept as a constant, the average loss does decrease but not converge to zero as p_1 increases.

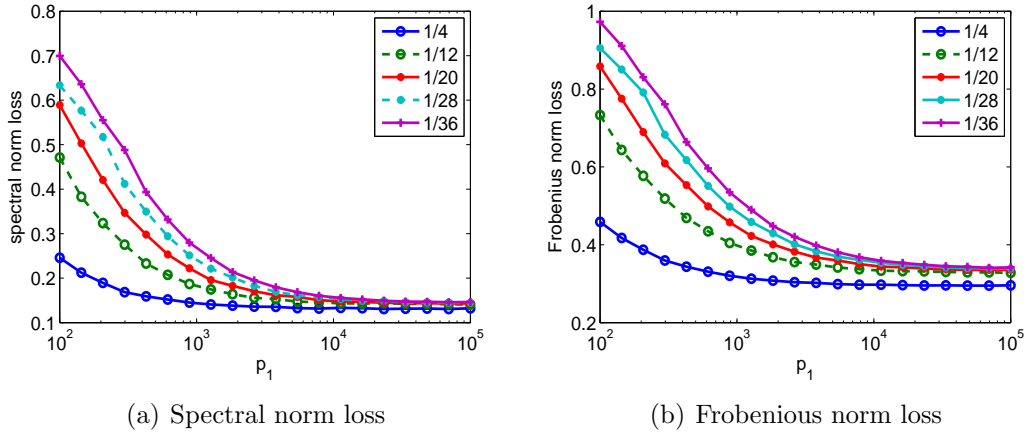


Figure A.2: Losses for settings with singular values of A being $\{j^{-1}, j = 1, 2, 3, \dots\}$, $p_2 = 1000$, $m_2 = 50$, $m_1/p_1 = 1/4, 1/12, 1/20, 1/28, 1/36$, and $p_1 = 100, \dots, 100,000$.

A.3.2 Technical Tools

We collect important technical tools in this section. The first lemma is about the inequalities of singular values in the perturbed matrix.

Lemma A.3.1. *Suppose $X \in \mathbb{R}^{p \times n}$, $Y \in \mathbb{R}^{p \times n}$, $\text{rank}(X) = a$, $\text{rank}(Y) = b$,*

1. $\sigma_{a+b+1-r}(X+Y) \leq \min(\sigma_{a+1-r}(X), \sigma_{b+1-r}(Y))$ for $r \geq 1$;
2. if we further have $X^\top Y = 0$, we must have $a+b \leq n$, $\sigma_r(X+Y) \geq \max(\sigma_r(X), \sigma_r(Y))$ for $r \geq 1$.

Lemma A.3.2. Suppose $X \in \mathbb{R}^{p \times n}, Y \in \mathbb{R}^{n \times m}$ are two arbitrary matrices, denote $\|\cdot\|_q, \|\cdot\|$ as the Schatten- q norm and spectral norm respectively, then we have

$$\|XY\|_q \leq \|X\|_q \cdot \|Y\|. \quad (\text{A.120})$$

The following two lemmas provide examples that illustrate NNM fails to recover \hat{A}_{22} .

Lemma A.3.3. Assume $A = B_1 B_2^T$, where $B_1 \in \mathbb{R}^{p_1 \times r}$ and $B_2 \in \mathbb{R}^{p_2 \times r}$ are two i.i.d. standard Gaussian matrices. Let A is divided into blocks as (3.1). Suppose

$$r \leq \frac{1}{400} \min(p_1, p_2), \quad m_1 \leq \frac{1}{25} p_1, \quad m_2 \leq \frac{1}{25} p_2, \quad (\text{A.121})$$

then the NNM (3.3) fails to recover A_{22} with probability at least $1 - 12 \exp(-\min(p_1, p_2)/400)$.

Lemma A.3.4. Denote $\mathbf{1}_p$ as the p -dimensional vector with all entries 1. Suppose $A = \mathbf{1}_{p_1} \cdot \mathbf{1}_{p_2}^T$, and A is divided into blocks as (3.1). Then the NNM (3.3) yields

$$\hat{A}_{22} = \min \left\{ \sqrt{\frac{m_1 m_2}{(p_1 - m_1)(p_2 - m_2)}}, 1 \right\} \mathbf{1}_{p_1 - m_1} \mathbf{1}_{p_2 - m_2}^T.$$

The following result is on the norm of a random submatrix of a given orthonormal matrix.

Lemma A.3.5. Suppose $U \in \mathbb{R}^{p \times d}$ is a fixed matrix with orthonormal columns (hence $d \leq p$). Denote $W = \max_{1 \leq i \leq p} \frac{p}{d} \cdot \sum_{j=1}^d u_{ij}^2$. Suppose we uniform randomly draw n rows (with or without replacement) from U and note the index as Ω and denote

$$U_\Omega = \begin{bmatrix} U_{\Omega(1)} \\ \vdots \\ U_{\Omega(n)} \end{bmatrix}.$$

When $n \geq \frac{4Wd(\log d+c)}{(1-\alpha)^2}$ for some $0 < \alpha < 1$ and $c > 1$, we have

$$\|\sigma_{\min}(U_{\Omega})\| \geq \sqrt{\frac{\alpha n}{p}}$$

with probability $1 - 2e^{-c}$.

The following results is about the spectral norm of the submatrix of a random orthonormal matrix.

Lemma A.3.6. *Suppose $U \in \mathbb{R}^{p \times d}$ ($d \leq p$) is with random orthonormal columns with Haar measure. For all $0 < \alpha_1 < 1 < \alpha_2$, there exists constant $C, \delta > 0$ depending only on α_1, α_2 such that when $p \geq n \geq \min\{Cd, p\}$, we have*

$$\sqrt{\frac{\alpha_1 n}{p}} \leq \sigma_{\min}(U_{[1:n,:]}) \leq \|U_{[1:n,:]}\| \leq \sqrt{\frac{\alpha_2 n}{p}} \quad (\text{A.122})$$

with probability at least $1 - \exp(-\delta n)$.

Proof of the Technical Lemmas

Proof of Lemma A.3.1.

1. First, by a well-known fact about best low-rank approximation,

$$\sigma_{a+b+1-r}(X+Y) = \min_{M \in \mathbb{R}^{p \times n}, \text{rank}(M) \leq a+b-r} \|X+Y-M\|.$$

Hence,

$$\sigma_{a+b+1-r}(X+Y) \leq \|X+Y - (X_{\max(a-r)} + Y)\| = \|X_{-\max(a-r)}\| = \sigma_{a+1-r}(X);$$

similarly $\sigma_{a+b+1-r}(X+Y) \leq \sigma_{b+1-r}(Y)$.

2. When we further have $X^\top Y = 0$, we know the column space of X and Y are orthogonal, then we have $\text{rank}(X + Y) = \text{rank}(X) + \text{rank}(Y) = a + b$, which means $a + b \leq n$. Next, note that

$$(X + Y)^\top(X + Y) = X^\top X + Y^\top Y + X^\top Y + Y^\top X = X^\top X + Y^\top Y,$$

if we note $\lambda_r(\cdot)$ as the r -th largest eigenvalue of the matrix, then we have

$$\begin{aligned} \sigma_r^2(X + Y) &= \lambda_r((X + Y)^\top(X + Y)) = \lambda_r(X^\top X + Y^\top Y) \\ &\geq \max(\lambda_r(X^\top X), \lambda_r(Y^\top Y)) = \max(\sigma_r^2(X), \sigma_r^2(Y)). \end{aligned}$$

□

Proof of Lemma A.3.2. Since

$$\|XY\|_q = \sqrt[q]{\sum_i \sigma_i^q(XY)}, \quad \|X\|_q = \sqrt[q]{\sum_i \sigma_i^q(X)},$$

it suffices to show $\sigma_i(XY) \leq \sigma_i(X)\|Y\|$. To this end, we have

$$\begin{aligned} \sigma_i(X) &= \min_{M \in \mathbb{R}^{p \times m}, \text{rank}(M) \leq i-1} \|XY - M\| \leq \|XY - X_{\max(i-1)}Y\| \\ &= \|X_{-\max(i-1)}Y\| \leq \sigma_i(X)\|Y\|, \end{aligned}$$

which finishes the proof of this lemma. □

Proof of Lemma A.3.3. Since B_1 and B_2 and their submatrices are all i.i.d. standard matrices, by the random matrix theory (Corollary 5.35 in Vershynin (2011)), for $t > 0$, we have with probability at least $1 - 12 \exp(-t^2/2)$, the following inequalities

hold,

$$\begin{aligned} \lambda_r(A) &\geq \lambda_{\min}(B_1)\lambda_{\min}(B_2) \geq (\sqrt{p_1} - \sqrt{r} - t)(\sqrt{p_2} - \sqrt{r} - t) \\ &\stackrel{(A.121)}{\geq} \left(\frac{19}{20}\sqrt{p_1} - t\right) \left(\frac{19}{20}\sqrt{p_2} - t\right) \end{aligned} \quad (A.123)$$

$$\begin{aligned} \|A_{1\bullet}\| &= \|B_{1,[1:m_1,:]}B_2^T\| \leq (\sqrt{m_1} + \sqrt{r} + t)(\sqrt{p_2} + \sqrt{r} + t) \\ &\stackrel{(A.121)}{\leq} \left(\frac{1}{4}\sqrt{p_1} + t\right) \left(\frac{21}{20}\sqrt{p_2} + t\right) \end{aligned} \quad (A.124)$$

and

$$\begin{aligned} \|A_{21}\| &= \|B_{1,[(m_1+1):p_1,:]}B_{2,[1:m_2,:]}^T\| \leq (\sqrt{p_1} + \sqrt{r} + t)(\sqrt{m_2} + \sqrt{r} + t) \\ &\stackrel{(A.121)}{\leq} \left(\frac{21}{20}\sqrt{p_1} + t\right) \left(\frac{1}{4}\sqrt{p_2} + t\right). \end{aligned} \quad (A.125)$$

Denote

$$A_0 = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & 0 \end{bmatrix}$$

and set $t = \frac{1}{20} \min(\sqrt{p_1}, \sqrt{p_2})$. Since $\|A_0\|_* \leq \|A_{1\bullet}\|_* + \|A_{21}\|_*$, we have

$$P\left(\|A\|_* \geq \frac{326}{400}\sqrt{p_1 p_2}\right) \geq 1 - 12 \exp(-\min(p_1, p_2)/400) \quad (A.126)$$

and

$$P\left(\|A_0\|_* \leq \frac{264}{400}\sqrt{p_1 p_2}\right) \geq 1 - 12 \exp(-\min(p_1, p_2)/400). \quad (A.127)$$

Hence, with probability at least $1 - 12 \exp(-\min(p_1, p_2)/400)$, $\|A_0\|_* < \|A\|_*$, which implies that the NNM (3.3) fails to recover A_{22} . \square

Proof of Lemma A.3.4. For convenience, we denote $x \wedge y = \min(x, y)$ for any two real numbers x, y . First, we can extend the unit vectors $\frac{1}{\sqrt{m_1}}\mathbf{1}_{m_1}$, $\frac{1}{\sqrt{m_2}}\mathbf{1}_{m_2}$, $\frac{1}{\sqrt{p_1-m_1}}\mathbf{1}_{p_1-m_1}$ and $\frac{1}{\sqrt{p_2-m_2}}\mathbf{1}_{p_2-m_2}$ into orthogonal matrices, which we denote as $U_{m_1} \in \mathbb{R}^{m_1 \times m_1}$,

$U_{m_2} \in \mathbb{R}^{m_2 \times m_2}$, $U_{p_1-m_1} \in \mathbb{R}^{(p_1-m_1) \times (p_1-m_1)}$, $U_{p_2-m_2} \in \mathbb{R}^{(p_2-m_2) \times (p_2-m_2)}$. Next, for all $A'_{22} \in \mathbb{R}^{(p_1-m_1) \times (p_2-m_2)}$, we must have

$$\begin{aligned} \left\| \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A'_{22} \end{bmatrix} \right\|_* &= \left\| \begin{bmatrix} U_{m_1}^\top & 0 \\ 0 & U_{p_1-m_1}^\top \end{bmatrix} \cdot \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A'_{22} \end{bmatrix} \cdot \begin{bmatrix} U_{m_2} & 0 \\ 0 & U_{p_2-m_2} \end{bmatrix} \right\|_* \\ &\triangleq \left\| \begin{bmatrix} E_{11} & E_{12} \\ E_{21} & U_{p_1-m_1}^\top A'_{22} U_{p_2-m_2} \end{bmatrix} \right\|_* , \end{aligned}$$

where $E_{11} \in \mathbb{R}^{m_1 \times m_2}$, $E_{12} \in \mathbb{R}^{m_1 \times (p_2-m_2)}$, $E_{21} \in \mathbb{R}^{(p_1-m_1) \times m_2}$ are with the first entry $\sqrt{m_1 m_2}$, $\sqrt{m_1(p_2-m_2)}$ and $\sqrt{m_2(p_1-m_1)}$ respectively and other entries 0. Therefore, we can see

$$\left\| \begin{bmatrix} E_{11} & E_{12} \\ E_{21} & U_{p_1-m_1}^\top A'_{22} U_{p_2-m_2} \end{bmatrix} \right\|_* \geq \left\| \begin{bmatrix} \sqrt{m_1 m_2} & \sqrt{m_1(p_2-m_2)} \\ \sqrt{m_2(p_1-m_1)} & [U_{p_1-m_1}^\top A'_{22} U_{p_2-m_2}]_{[1,1]} \end{bmatrix} \right\|_*$$

and the equality holds if and only if $U_{p_1-m_1}^\top A'_{22} U_{p_2-m_2}$ is zero except the first entry.

By some calculation, we can see the nuclear norm of 2-by-2 matrix

$$\left\| \begin{bmatrix} \sqrt{m_1 m_2} & \sqrt{m_1(p_2-m_2)} \\ \sqrt{m_2(p_1-m_1)} & x \end{bmatrix} \right\|_*$$

achieves its minimum if and only if

$$x = \sqrt{m_1 m_2} \wedge \sqrt{(p_1-m_1)(p_2-m_2)}.$$

Hence, A'_{22} achieves the minimum of $\left\| \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A'_{22} \end{bmatrix} \right\|_*$ if and only if

$$U_{p_1-m_1}^\top A'_{22} U_{p_2-m_2} = \begin{bmatrix} \sqrt{m_1 m_2} \wedge \sqrt{(p_1-m_1)(p_2-m_2)} & 0 & \cdots \\ 0 & 0 & \\ \vdots & & \ddots \end{bmatrix},$$

which means the minimizer $A'_{22} = \left(\sqrt{\frac{m_1 m_2}{(p_1-m_1)(p_2-m_2)}} \wedge 1 \right) \cdot \mathbf{1}_{p_1-m_1} \mathbf{1}_{p_2-m_2}^\top$. \square

Proof of Lemma A.3.5. The proof of this lemma relies on operator-Bernstein's inequality for sampling (Theorem 1 in Gross and Nesme (2010)). For two symmetric matrices A, B , we say $A \preceq B$ if $B - A$ is positive definite. By assumption, $\{U_{\Omega(j)\bullet}, j = 1, \dots, n\}$ are uniformly random samples (with or without replacement) from $\{U_{i\bullet}, i = 1, \dots, n\}$. Suppose

$$X_i = U_{i\bullet}^\top U_{i\bullet} - \frac{1}{p} I_d, \quad i = 1, \dots, p, \quad (\text{A.128})$$

then X_i are symmetric matrices, $X_{\Omega(j)}, j = 1, \dots, n$ are uniformly random samples (with or without replacement) from $\{X_1, \dots, X_p\}$. In addition, we have

$$EX_j = \frac{1}{p} \sum_{i=1}^p U_{i\bullet}^\top U_{i\bullet} - \frac{1}{p} I_d = \frac{1}{p} U^\top U - \frac{1}{p} I_d = 0$$

$$\|X_j\| \leq \max_{1 \leq i \leq p} \left\| U_{i\bullet}^\top U_{i\bullet} - \frac{1}{p} I_d \right\| \leq \max_{1 \leq i \leq p} \max \left\{ \|U_{i\bullet}^\top U_{i\bullet}\|, \frac{1}{p} \|I_d\| \right\} \leq \frac{Wd}{p}$$

$$\begin{aligned} EX_j^2 &= \frac{1}{p} \sum_{i=1}^p \left(U_{i\bullet}^\top U_{i\bullet} - \frac{1}{p} I_d \right)^2 = \frac{1}{p} \sum_{i=1}^p \left(U_{i\bullet}^\top U_{i\bullet} U_{i\bullet}^\top U_{i\bullet} - \frac{2}{p} U_{i\bullet}^\top U_{i\bullet} + \frac{1}{p^2} I_d \right) \\ &= \frac{1}{p} \sum_{i=1}^p \|U_{i\bullet}\|_2^2 \cdot U_{i\bullet}^\top U_{i\bullet} - \frac{1}{p^2} I_d \\ &\preceq \frac{1}{p} \cdot \frac{Wd}{p} \sum_{i=1}^p U_{i\bullet}^\top U_{i\bullet} - \frac{1}{p^2} I_d \preceq \frac{Wd-1}{p^2} I_d \end{aligned}$$

For all $0 < \alpha < 1$, by Theorem 1 in Gross and Nesme (2010),

$$\begin{aligned}
P\left(\|U_\Omega\| \leq \sqrt{\frac{\alpha n}{p}}\right) &= P\left(U_\Omega^\top U_\Omega \preceq \frac{\alpha n}{p} I_d\right) = P\left(\sum_{j=1}^n U_{\Omega(j)}^\top \bullet U_{\Omega(j)} \bullet \preceq \frac{\alpha n}{p} I_d\right) \\
&= P\left(\sum_{j=1}^n X_j \preceq -\frac{(1-\alpha)n}{p} I_d\right) \leq P\left(\left\|\sum_{j=1}^n X_j\right\| \geq \frac{(1-\alpha)n}{p}\right) \\
&\leq 2d \exp\left(-\min\left(\frac{((1-\alpha)n/p)^2}{4n(Wd-1)/p^2}, \frac{(1-\alpha)n/p}{2Wd/p}\right)\right) \\
&\leq 2d \exp\left(-\frac{n(1-\alpha)^2}{4Wd}\right) \leq 2 \exp(-c).
\end{aligned}$$

The last inequality is due to the assumption that

$$n \geq \frac{4Wd(\log d + c)}{(1-\alpha)^2}.$$

□

Proof of Lemma A.3.6. By the assumption on n , we have $n \geq p$ or $n \geq Cd$. When $n \geq p$, we know $n = p$ and $U_{[1:n,:]} = U$ is an orthogonal matrix, which means (A.122) is clearly true. Hence, we only need to prove the theorem under the assumption that $p \geq n$ is true. In this case, we must have $n \geq Cd$.

Since U has random orthonormal columns with Haar measure, for any fixed vector $v \in \mathbb{R}^d$, Uv is identical distributed as

$$\|x\|_2^{-1} (x_1, x_2, \dots, x_p), \quad \text{where } x_1, \dots, x_p \stackrel{iid}{\sim} N(0, 1)$$

Hence, $U_{[1:n,:]}v$ is identical distributed with $\|x\|_2^{-1}(x_1, \dots, x_n)$ and

$$\|U_{[1:n,:]}v\|_2 \text{ is identical distributed as } \sqrt{\left(\sum_{i=1}^n x_i^2\right)\left(\sum_{i=1}^p x_i^2\right)^{-1}}, \quad (\text{A.129})$$

which is the also the square root of Beta distribution. Denote

$$\alpha'_1 = \frac{1 + \alpha_1}{2}, \quad \alpha'_2 = \frac{1 + \alpha_2}{2}. \quad (\text{A.130})$$

By Lemma 1 in Laurent and Massart (2000), when x_1, \dots, x_p are i.i.d. standard normal, we have

$$1 - 2\sqrt{C'} \leq \frac{\sum_{i=1}^n x_i^2}{n} \leq 1 + 2\sqrt{C'} + 2C'$$

$$1 - 2\sqrt{\frac{C'n}{p}} \leq \frac{\sum_{i=1}^p x_i^2}{p} \leq 1 + 2\sqrt{\frac{C'n}{p}} + \frac{2C'n}{p}$$

both hold with probability at least $1 - 4\exp(-C'n)$. Here we let $C' > 0$ be small enough and only depending on α_1, α_2 such that

$$\alpha'_1 \leq \frac{1 - 2\sqrt{C'}}{1 + 2\sqrt{C'} + 2C'}, \quad \frac{1 + 2\sqrt{C'} + 2C'}{1 - 2\sqrt{C'}} \leq \alpha'_2.$$

Combining the previous inequalities and (A.129), we have for any fixed unit vector $v \in \mathbb{R}^d$,

$$\frac{\alpha'_1 n}{p} \leq \|U_{[1:n, \cdot]} v\|_2^2 \leq \frac{\alpha'_2 n}{p} \quad (\text{A.131})$$

with probability at least $1 - 4\exp(-C'n)$, where C' only depends on α'_1, α'_2 . Next, based on Lemma 2.5 in Vershynin (2011), we can construct an ε -net on the unit sphere of \mathbb{R}^d as B , such that $|B| \leq (1 + 2/\varepsilon)^d$, where $\varepsilon > 0$ is to be determined later. Under the event that $\{\forall v \in B, (\text{A.131}) \text{ holds}\}$, we suppose

$$\kappa_1 = \min_{\|v\|_2=1} \|U_{[1:n, \cdot]} v\|_2^2, \quad \kappa_2 = \max_{\|v\|_2=1} \|U_{[1:n, \cdot]} v\|_2^2.$$

For any v in the unit sphere of \mathbb{R}^d , there must exists $v' \in B$ such that $\|v - v'\|_2 \leq \varepsilon$,

which yields,

$$\|U_{[1:n,:]}v\|_2 \leq \|U_{[1:n,:]}v'\|_2 + \|U_{[1:n,:]}(v - v')\|_2 \leq \sqrt{\alpha'_2 n/p} + \kappa_2 \varepsilon$$

$$\|U_{[1:n,:]}v\|_2 \geq \|U_{[1:n,:]}v'\|_2 - \|U_{[1:n,:]}(v - v')\|_2 \geq \sqrt{\alpha'_1 n/p} - \varepsilon \kappa_2$$

These implies that $\kappa_2 \leq \sqrt{\alpha'_2 n/p}/(1 - \varepsilon)$, $\kappa_1 \geq \sqrt{\alpha'_1 n/p} - \varepsilon \kappa_2 \geq \sqrt{\alpha'_1 n/p} - \sqrt{\alpha'_2 n/p} \cdot \varepsilon/(1 - \varepsilon)$. Hence, we can take ε depending on α_1, α_2 such that $\kappa_2 \leq \sqrt{\alpha_2 n/p}$, $\kappa_1 \geq \sqrt{\alpha_1 n/p}$, which implies (A.122).

Finally we estimate the probability that the event $\{\forall v \in B, \text{(A.131) holds}\}$ happens. We choose $C \geq 4d \log(1 + 2/\varepsilon)/C'$ that only depends on α_1 and α_2 . If $n \geq Cd$,

$$C'n/2 \geq d \log(1 + 2/\varepsilon) + \log 4.$$

so

$$1 - (1 + 2/\varepsilon)^d \cdot 4 \exp(-C'n) = 1 - \exp(d \log(1 + 2/\varepsilon) + \log 4 - C'n) \geq 1 - \exp(-nC'/2)$$

Finally, we finish the proof of the lemma by setting $\delta = C'/2$. \square

A.3.3 Proofs of the Results in the Main Paper

We prove Proposition 3.2.1, Theorems 3.3.1 and 3.3.2, Lemma A.3.7, Lemma A.3.8, Theorem 3.3.3, Corollary 3.3.1 and Corollary 3.3.2 in this section.

Proof of Proposition 3.2.1

Since $A_{1\bullet}$ is of rank r , which is the same as A , all rows of A must be linear combinations of the rows of $A_{1\bullet}$. This implies all rows of $A_{\bullet 1}$ is a linear combination of A_{11} . Since $\text{rank}(A_{\bullet 1}) = r$, we must have $\text{rank}(A_{11}) \geq r$. Besides, $\text{rank}(A_{11}) \leq \text{rank}(A) = r$

since A_{11} is a submatrix of A . So $\text{rank}(A_{11}) = r$. Similarly, rows of $A_{\bullet 1}$ is the linear combination of A_{11} , so we have

$$A_{21} = A_{21}P_{A_{11}} = A_{21}A_{11}^\top(A_{11}A_{11}^\top)^\dagger A_{11} = A_{21}V\Sigma U^\top(U\Sigma^2U^\top)^\dagger A_{11} = (A_{21}V\Sigma^{-1}U^\top)A_{11},$$

namely rows of A_{21} is a linear combination of A_{11} . By the argument before, we know A_{22} can be represented as the same linear combination of A_{12} as A_{21} by A_{11} , so we have $A_{22} = (A_{21}V\Sigma^{-1}U^\top)A_{12} = A_{21}V\Sigma^{-1}U^\top A_{12} = A_{21}A_{11}^\dagger A_{12}$, which concludes the proof. \square

Proof of Theorem 3.3.1

Suppose $M \in \mathbb{R}^{m_1 \times r}$, $N \in \mathbb{R}^{m_2 \times r}$ are column orthonormalized matrices of U_{11} and V_{11} . $\hat{M} \in \mathbb{R}^{m_1 \times r}$ and $\hat{N} \in \mathbb{R}^{m_2 \times r}$ are the first r left singular vectors of $A_{1\bullet}$ and $A_{\bullet 1}$, respectively. Also, recall that we use $P_U = U(U^\top U)^\dagger U^\top$ to represent the projection onto the column space of U .

1. We first give the lower bound for $\sigma_{\min}(\hat{M}^\top M)$, $\sigma_{\min}(\hat{N}^\top N)$ by the unilateral perturbation bound result in Cai and Zhang (2014a). Since,

$$P_{U_{11}}A_{1\bullet} = P_{U_{11}}U_{1\bullet}\Sigma V^\top = [U_{11}\Sigma_1, P_{U_{11}}U_{12}\Sigma_2]V^\top,$$

$$P_{U_{11}^\perp}A_{1\bullet} = P_{U_{11}^\perp}U_{1\bullet}\Sigma V^\top = [0, P_{U_{11}^\perp}U_{12}\Sigma_2]V^\top,$$

by V is an orthogonal matrix, we can see

$$\sigma_r(P_{U_{11}}A_{1\bullet}) = \sigma_r([U_{11}\Sigma_1 \quad P_{U_{11}}U_{12}\Sigma_2]) \geq \sigma_r(U_{11}\Sigma_1) \geq \sigma_r(A)\sigma_{\min}(U_{11}),$$

$$\|P_{U_{11}^\perp}A_{1\bullet}\| = \|P_{U_{11}^\perp}U_{12}\Sigma_2\| \leq \|P_{U_{11}^\perp}U_{12}\|\|\Sigma_2\| \leq \sigma_{r+1}(A).$$

So $\sigma_r(P_{U_{11}}A_{1\bullet}) \geq \|P_{U_{11}^\perp}A_{1\bullet}\|$. Besides, $\text{rank}(P_{U_{11}}A_{1\bullet}) \leq r$. Apply the unilateral perturbation bound result in Cai and Zhang (2014a) by setting $X = P_{U_{11}}A_{1\bullet}$, $Y = P_{U_{11}^\perp}A_{1\bullet}$, we have

$$\sigma_{\min}^2(\hat{M}^\top M) \leq 1 - \left(\frac{\|Y \cdot P_{X^\top}\| \cdot \sigma_{r+1}(A)}{\sigma_r^2(A)\sigma_{\min}^2(U_{11}) - \sigma_{r+1}^2(A)} \right)^2. \quad (\text{A.132})$$

Moreover, $A_{1\bullet} = [U_{11} \ U_{12}] \text{diag}(\Sigma_1, \Sigma_2) V^\top = [U_{11}\Sigma_1 \ U_{12}\Sigma_2] V^\top$, and hence,

$$\begin{aligned} \|Y P_{X^\top}\| &= \left\| P_{U_{11}^\perp} A_{1\bullet} \cdot P_{(P_{U_{11}}A_{1\bullet})^\top} \right\| = \left\| [0 \ P_{U_{11}^\perp} U_{12}\Sigma_2] V^\top \cdot P_{V \cdot [U_{11}\Sigma_1 \ P_{U_{11}}U_{12}\Sigma_2]^\top} \right\| \\ &= \left\| [0 \ P_{U_{11}^\perp} U_{12}\Sigma_2] \cdot P_{[U_{11}\Sigma_1 \ P_{U_{11}}U_{12}\Sigma_2]^\top} \right\| \\ &= \sup_{x \in \mathbb{R}^{p_2}, \|x\|_2=1} [0 \ P_{U_{11}^\perp} U_{12}\Sigma_2] \cdot P_{[U_{11}\Sigma_1 \ P_{U_{11}}U_{12}\Sigma_2]^\top} x. \end{aligned}$$

When $\|x\|_2 = 1$, let y denote the projection of x onto the column space of $[U_{11}\Sigma_1 \ P_{U_{11}}U_{12}\Sigma_2]^\top$. Then $\|y\|_2 \leq 1$ and y is in the column space of $[U_{11}\Sigma_1 \ P_{U_{11}}U_{12}\Sigma_2]^\top$. Hence,

$$\frac{\|y_{[1:m_1]}\|_2}{\|y_{[(m_1+1):p_1]}\|_2} \geq \frac{\sigma_{\min}(U_{11}\Sigma_1)}{\|P_{U_{11}}U_{12}\Sigma_2\|} \geq \frac{\sigma_{\min}(U_{11})\sigma_r(A)}{\sigma_{r+1}(A)},$$

$$\text{and } \|y_{[(m_1+1):p_1]}\|_2^2 + \|y_{[1:m_1]}\|_2^2 \leq 1,$$

which implies $\|y_{[(m_1+1):p_1]}\|_2^2 \leq \sigma_{r+1}^2(A)/\sigma_{\min}^2(U_{11})\sigma_r^2(A) + \sigma_{r+1}^2(A)$. Hence for all $x \in \mathbb{R}^{p_2}$ such that $\|x\|_2 = 1$,

$$\begin{aligned} \left\| [0 \ P_{U_{11}^\perp} U_{12}\Sigma_2] \cdot P_{[U_{11}\Sigma_1 \ P_{U_{11}}U_{12}\Sigma_2]^\top} x \right\| &\leq \|P_{U_{11}^\perp} U_{12}\Sigma_2\| \cdot \|y_{[m_1+1:p_1]}\|_2 \\ &\leq \sigma_{r+1}(A) \frac{\sigma_{r+1}(A)}{\sqrt{\sigma_{r+1}^2(A) + \sigma_{\min}^2(U_{11})\sigma_r^2(A)}}. \end{aligned}$$

This yields

$$\|YP_{X^\tau}\| = \|P_{U_{11}^\perp} A_{1\bullet} \cdot P_{(P_{U_{11}} A_{1\bullet})}\| \leq \sigma_{r+1}^2(A) / \sqrt{\sigma_{r+1}^2(A) + \sigma_{\min}^2(U_{11})\sigma_r^2(A)}.$$

Combining (A.132), we have

$$\begin{aligned} & \sigma_{\min}^2(\hat{M}^\top M) \\ & \geq 1 - \left(\frac{\sigma_{r+1}^3(A)}{\sqrt{\sigma_{r+1}^2(A) + \sigma_{\min}^2(U_{11})\sigma_r^2(A)} (\sigma_r^2(A)\sigma_{\min}^2(U_{11}) - \sigma_{r+1}^2(A))} \right)^2. \end{aligned} \quad (\text{A.133})$$

Since $\sigma_{\min}(U_{11})\sigma_r(A) \geq 2\sigma_{r+1}(A)$, we have

$$\sigma_{\min}^2(\hat{M}^\top M) \geq 1 - \left(\frac{1}{\sqrt{5} \cdot 3} \right)^2 \geq \frac{44}{45}.$$

Similarly, we also have $\sigma_{\min}^2(\hat{N}^\top N) \geq \frac{44}{45}$.

2. Following by (3.8),

$$\begin{aligned} \hat{A}_{22} &= U_{2\bullet} \Sigma V_{1\bullet}^\top \hat{N} \left(\hat{M}^\top (U_{1\bullet} \Sigma V_{1\bullet}^\top) \hat{N} \right)^{-1} \hat{M}^\top U_{1\bullet} \Sigma V_{2\bullet}^\top \\ &= \left(U_{21} \Sigma_1 V_{11}^\top \hat{N} + U_{22} \Sigma_2 V_{12}^\top \hat{N} \right) \left(\hat{M}^\top U_{11} \Sigma_1 V_{11}^\top \hat{N} + \hat{M}^\top U_{12} \Sigma_2 V_{12}^\top \hat{N} \right)^{-1} \\ &\quad \cdot \left(\hat{M}^\top U_{11} \Sigma_1 V_{21}^\top + \hat{M}^\top U_{12} \Sigma_2 V_{22}^\top \right). \end{aligned}$$

Let ‘‘L’’, ‘‘M’’, ‘‘R’’ stand for ‘‘Left’’, ‘‘Middle’’ and ‘‘Right’’,

$$B_L = U_{21} \Sigma_1 V_{11}^\top \hat{N}, \quad E_L = U_{22} \Sigma_2 V_{12}^\top \hat{N}; \quad (\text{A.134})$$

$$B_M = \hat{M}^\top U_{11} \Sigma_1 V_{11}^\top \hat{N}, \quad E_M = \hat{M}^\top U_{12} \Sigma_2 V_{12}^\top \hat{N}; \quad (\text{A.135})$$

$$B_R = \hat{M}^\top U_{11} \Sigma_1 V_{21}^\top, \quad E_R = \hat{M}^\top U_{12} \Sigma_2 V_{22}^\top. \quad (\text{A.136})$$

By Lemma A.3.2 in the Supplement, we can see the following properties of these

matrices,

$$\|E_L\| \leq \sigma_{r+1}(A), \quad \|E_M\| \leq \sigma_{r+1}(A), \quad \|E_R\| \leq \sigma_{r+1}(A), \quad (\text{A.137})$$

$$\|E_L\|_q \leq \|\Sigma_2\|_q, \quad \|E_M\|_q \leq \|\Sigma_2\|_q, \quad \|E_R\|_q \leq \|\Sigma_2\|_q, \quad (\text{A.138})$$

$$\begin{aligned} \sigma_{\min}(B_M) &= \sigma_{\min} \left(\hat{M}^\top (P_M U_{11}) \Sigma_1 (V_{11}^\top P_N) \hat{N} \right) \\ &= \sigma_{\min} \left((\hat{M}^\top M) (M^\top U_{11}) \Sigma_1 (V_{11}^\top N) (N^\top \hat{N}) \right) \\ &\geq \sigma_{\min}(\Sigma_1) \sigma_{\min}(U_{11}) \sigma_{\min}(V_{11}) \sigma_{\min}(\hat{M}^\top M) \sigma_{\min}(\hat{N}^\top N) \\ &\geq \frac{44}{45} \sigma_r(A) \sigma_{\min}(U_{11}) \sigma_{\min}(V_{11}), \end{aligned} \quad (\text{A.139})$$

$$\|B_M^{-1}\| = \sigma_{\min}^{-1}(B_M) \leq \frac{45}{44 \sigma_r(A) \sigma_{\min}(U_{11}) \sigma_{\min}(V_{11})}, \quad (\text{A.140})$$

$$\hat{A}_{22} = (B_L + E_L)(B_M + E_M)^{-1}(B_R + E_R), \quad B_L B_M^{-1} B_R = U_{21} \Sigma_1 V_{21}^\top, \quad (\text{A.141})$$

$$\begin{aligned} \|B_L B_M^{-1}\| &= \|U_{21} \Sigma_1 (V_{11}^\top \hat{N}) (V_{11}^\top \hat{N})^{-1} \Sigma^{-1} (\hat{M}^\top U_{11})^{-1}\| = \|U_{21} (\hat{M}^\top U_{11})^{-1}\| \\ &\leq \|(\hat{M}^\top M M^\top U_{11})^{-1}\| \leq \frac{1}{\sigma_{\min}(M^\top U_{11}) \sigma_{\min}(\hat{M}^\top M)} \leq \frac{\sqrt{45/44}}{\sigma_{\min}(U_{11})}, \end{aligned} \quad (\text{A.142})$$

$$\|B_M^{-1} B_R\| = \|(V_{11} \hat{N})^{-1} V_{21}^\top\| \leq \frac{\sqrt{45/44}}{\sigma_{\min}(V_{11})}. \quad (\text{A.143})$$

By (A.137), (A.139) and the assumption (3.10), we can see $\sigma_{\min}(B_M) > \|E_M\|$,

so

$$\hat{A}_{22} \stackrel{(\text{A.141})}{=} (B_L + E_L)(B_M^{-1} - B_M^{-1} E_M B_M^{-1} + B_M^{-1} E_M B_M^{-1} E_M B_M^{-1} - \dots)(B_R + E_R);$$

$$\begin{aligned}
& \|\hat{A}_{22} - B_L B_M^{-1} B_R\|_q \\
& \leq \|B_L B_M^{-1} E_M \sum_{i=0}^{\infty} (-B_M^{-1} E_M)^i B_M^{-1} B_R\|_q + \|E_L \sum_{i=0}^{\infty} (-B_M^{-1} E_M)^i B_M^{-1} B_R\|_q \\
& \quad + \|B_L B_M^{-1} \sum_{i=0}^{\infty} (-E_M B_M^{-1})^i E_R\|_q + \|E_L B_M^{-1} \sum_{i=0}^{\infty} (-E_M B_M^{-1})^i E_R\|_q \\
& \leq \|B_L B_M^{-1}\| \|E_M\|_q \sum_{i=0}^{\infty} \|E_M\|^i \|B_M^{-1}\|^i \|B_M^{-1} B_R\| \\
& \quad + \|E_L\|_q \sum_{i=0}^{\infty} \|B_M^{-1}\|^i \|E_M\|^i \|B_M^{-1} B_R\| \\
& \quad + \|B_L B_M^{-1}\| \sum_{i=0}^{\infty} \|E_M\|^i \|B_M^{-1}\|^i \|E_R\|_q \\
& \quad + \|E_L\| \sum_{i=0}^{\infty} \|B_M^{-1}\|^{i+1} \|E_M\|^i \|E_R\|_q \\
& \stackrel{(A.137)(A.138)}{\leq} \frac{\|B_L B_M^{-1}\| \|B_M^{-1} B_R\| + \|B_M^{-1} B_R\| + \|B_L B_M^{-1}\| + \|B_M^{-1}\| \sigma_{r+1}(A)}{1 - \sigma_{r+1}(A) \|B_M^{-1}\|} \|\Sigma_2\|_q \\
& \stackrel{(A.142)(A.143)}{\leq} \frac{1}{1 - \sigma_{r+1}(A) \|B_M^{-1}\|} \\
& \quad \cdot \left(\frac{45/44}{\sigma_{\min}(U_{11}) \sigma_{\min}(V_{11})} + \frac{\sqrt{45/44}}{\sigma_{\min}(U_{11})} + \frac{\sqrt{45/44}}{\sigma_{\min}(V_{11})} + \frac{45}{88} \right) \|\Sigma_2\|_q \\
& \leq \frac{\|A_{-\max(r)}\|_q}{1 - \frac{45\sigma_{r+1}(A)}{44\sigma_r(A) \sigma_{\min}(U_{11}) \sigma_{\min}(V_{11})}} \\
& \quad \cdot \left(\frac{45/44}{\sigma_{\min}(U_{11}) \sigma_{\min}(V_{11})} + \frac{\sqrt{45/44}}{\sigma_{\min}(U_{11})} + \frac{\sqrt{45/44}}{\sigma_{\min}(V_{11})} + \frac{45}{88} \right) \\
& \leq \frac{88}{43} \|A_{-\max(r)}\|_q \left(\frac{45/44}{\sigma_{\min}(U_{11}) \sigma_{\min}(V_{11})} + \frac{\sqrt{45/44}}{\sigma_{\min}(U_{11})} + \frac{\sqrt{45/44}}{\sigma_{\min}(V_{11})} + \frac{45}{88} \right).
\end{aligned}$$

Finally, since $A_{22} = U_{21}\Sigma_1V_{21}^\top + U_{22}\Sigma_2V_{22}^\top \stackrel{(A.141)}{=} B_L B_M^{-1} B_R + U_{22}\Sigma_2V_{22}^\top$, we have

$$\begin{aligned} \|\hat{A}_{22} - A_{22}\|_q &\leq \|\hat{A}_{22} - B_L B_M^{-1} B_R\|_q + \|U_{22}\Sigma_2V_{22}^\top\|_q \\ &\leq 3\|A_{-\max(r)}\|_q \left(1 + \frac{1}{\sigma_{\min}(U_{11})}\right) \left(1 + \frac{1}{\sigma_{\min}(V_{11})}\right). \quad \square \end{aligned}$$

Proof of Theorem 3.3.2

We only present proof for row thresholding as the column thresholding is essentially the same by working with A^T . Suppose M, N are orthonormal basis of column vectors of U_{11}, V_{11} . We denote $U_{[:,1:r]}^{(1)} = \hat{M}$, $V_{[:,1:r]}^{(2)} = \hat{N}$, which are exactly the same as the \hat{M} and \hat{N} in Algorithm 1. Similarly to the proof of Theorem 3.3.1, we have (A.133). Due to the assumption that $\sigma_r(A)\sigma_{\min}(U_{11})\sigma_{\min}(V_{11}) \geq 4\sigma_{r+1}(A)$, (A.133) yields

$$\sigma_{\min}^2(\hat{M}^\top M) \geq 3824/3825, \quad \sigma_{\min}^2(\hat{N}^\top N) \geq 3824/3825. \quad (\text{A.144})$$

As shown in the Supplementary material, we have

Lemma A.3.7. *Under the assumption of Theorem 3.3.2, we have $\hat{r} \geq r$.*

We next show (3.13) with the condition that $\hat{r} \geq r$ in steps.

1. Note that $A_{11} = U_{11}\Sigma_1V_{11}^\top + U_{12}\Sigma_2V_{12}^\top$, we consider the decompositions of Z and let

$$Z_{11} = U^{(2)\top}U_{11}\Sigma_1V_{11}^\top V^{(1)} + U^{(2)\top}U_{12}\Sigma_2V_{12}^\top V^{(1)},$$

$$Z_{11,[:,1:\hat{r}]} = U_{[:,1:\hat{r}]}^{(2)\top}U_{11}\Sigma_1V_{11}^\top V_{[:,1:\hat{r}]}^{(1)} + U_{[:,1:\hat{r}]}^{(2)\top}U_{12}\Sigma_2V_{12}^\top V_{[:,1:\hat{r}]}^{(1)} \triangleq B_{M,\hat{r}} + E_{M,\hat{r}}, \quad (\text{A.145})$$

$$Z_{21,[:,1:\hat{r}]} = U_{21}\Sigma_1V_{11}^\top V_{[:,1:\hat{r}]}^{(1)} + U_{22}\Sigma_2V_{12}^\top V_{[:,1:\hat{r}]}^{(1)} \triangleq B_{L,\hat{r}} + E_{L,\hat{r}}, \quad (\text{A.146})$$

$$Z_{12,[:,1:\hat{r},:]} = U_{[:,1:\hat{r}]}^{(2)\top}U_{11}\Sigma_1V_{21}^\top + U_{[:,1:\hat{r}]}^{(2)\top}U_{12}\Sigma_2V_{22}^\top \triangleq B_{R,\hat{r}} + E_{R,\hat{r}}. \quad (\text{A.147})$$

Note that the square matrix $U_{[:,1:r]}^{(2)\top}M \in \mathbb{R}^{r \times r}$ is a submatrix of $U_{[:,1:\hat{r}]}^{(2)\top}M \in \mathbb{R}^{\hat{r} \times r}$, we know

$$\sigma_{\min}(U_{[:,1:\hat{r}]}^{(2)\top}M) \geq \sigma_{\min}(U_{[:,1:r]}^{(2)\top}M) = \sigma_{\min}(\hat{M}M) \stackrel{(A.144)}{\geq} \sqrt{\frac{3824}{3825}}. \quad (\text{A.148})$$

Similarly, $\sigma_{\min}(V_{[:,1:\hat{r}]}^{(1)\top}N) \geq \sqrt{\frac{3824}{3825}}$. By M, N are the orthonormal basis of column vectors of U_{11}, V_{11} , we have $P_M = MM^\top, P_N = NN^\top$, and

$$\sigma_{\min}(U_{[:,1:\hat{r}]}^{(2)\top}U_{11}) \geq \sigma_{\min}(U_{[:,1:r]}^{(2)\top}M)\sigma_{\min}(M^\top U_{11}) \geq \sqrt{\frac{3824}{3825}}\sigma_{\min}(U_{11}); \quad (\text{A.149})$$

similarly, we also have

$$\sigma_{\min}(V_{[:,1:\hat{r}]}^{(1)\top}V_{11}) \geq \sqrt{\frac{3824}{3825}}\sigma_{\min}(V_{11}). \quad (\text{A.150})$$

(A.149) and (A.150) immediately yield

$$\sigma_r(B_{M,\hat{r}}) \geq \frac{3824}{3825}\sigma_{\min}(U_{11})\sigma_{\min}(\Sigma_1)\sigma_{\min}(V_{11}) = \frac{3824}{3825}\sigma_r(A)\sigma_{\min}(U_{11})\sigma_{\min}(V_{11}). \quad (\text{A.151})$$

Besides, we also have

$$\|E_{M,\hat{r}}\| \stackrel{(A.145)}{\leq} \|\Sigma_2\| = \sigma_{r+1}(A) \quad (\text{A.152})$$

2. Next, we consider the SVD of $Z_{11,[1:\hat{r},1:\hat{r}]}$

$$Z_{11,[1:\hat{r},1:\hat{r}]} = J\Lambda K^\top, \quad J, \Lambda, K \in \mathbb{R}^{\hat{r} \times \hat{r}}. \quad (\text{A.153})$$

For convenience, we denote $\Lambda_1 = \Lambda_{[1:r,1:r]}$, $\Lambda_2 = \Lambda_{[(r+1):\hat{r},(r+1):\hat{r}]}$,

$$J_1 = J_{[:,1:r]}, \quad J_2 = J_{[:,(r+1):\hat{r}]}, \quad K_1 = K_{[:,1:r]}, \quad K_2 = K_{[:,(r+1):\hat{r}]}, \quad (\text{A.154})$$

Suppose $M_Z \in \mathbb{R}^{\hat{r} \times r}$ is an orthonormal basis of the column space of $B_{M,\hat{r}}$; $N_Z \in \mathbb{R}^{\hat{r} \times r}$ is an orthonormal basis of the column space of $B_{M,\hat{r}}^\top$. Denote $\text{span}(\cdot)$ as the linear span of the column space of the matrix. We want to show $\text{span}(M_Z)$ is close to $\text{span}(J_1)$; while $\text{span}(N_Z)$ is close to $\text{span}(K_1)$. So in the rest of this step, we try to establish bounds for $\sigma_{\min}(J_1^\top M_Z)$ and $\sigma_{\min}(K_1^\top N_Z)$. Actually,

$$Z_{11,[1:\hat{r},1:\hat{r}]} = B_{M,\hat{r}} + E_{M,\hat{r}} = (B_{M,\hat{r}} + P_{M_Z} E_{M,\hat{r}}) + P_{M_Z^\perp} E_{M,\hat{r}}.$$

Now we set $X = (B_{M,\hat{r}} + P_{M_Z} E_{M,\hat{r}})$, $Y = P_{M_Z^\perp} E_{M,\hat{r}}$, then we have

$$\begin{aligned} \sigma_r(X) &\geq \sigma_r(B_{M,\hat{r}}) - \|P_{M_Z} E_{M,\hat{r}}\| \stackrel{(\text{A.151})}{\geq} \frac{3824}{3825} \sigma_r(A) \sigma_{\min}(U_{11}) \sigma_{\min}(V_{11}) - \sigma_{r+1}(A), \\ &\stackrel{(3.12)}{\geq} \sigma_{r+1}(A) \stackrel{(\text{A.152})}{\geq} \|E_{M,\hat{r}}\| \geq \|Y\|. \end{aligned}$$

Besides, by the definition of $B_{M,\hat{r}}$ and M_Z we know $\text{rank}(X) \leq r$. Also based on the definition of Y , we know $P_X Y = 0$. Now the unilateral perturbation bound in Cai and Zhang (2014a) yields

$$\sigma_{\min}^2(M_Z^\top J_1) \geq 1 - \left(\frac{\sigma_r(X) \cdot \|Y\|}{\sigma_r^2(X) - \|Y\|^2} \right)^2. \quad (\text{A.155})$$

The right hand side of the inequality above is an increasing function of $\sigma_r(X)$. Since $\sigma_r(X) \geq \frac{3824}{3825} \sigma_r(A) \sigma_{\min}(U_{11}) \sigma_{\min}(V_{11}) - \sigma_{r+1}(A) \geq (3 - \frac{4}{3825}) \sigma_{r+1}(A) \geq (3 - \frac{4}{3825}) \|Y\|$,

$$\sigma_{\min}^2(J_1^\top M_Z) \geq 1 - \left(\frac{3 - 4/3825}{(3 - 4/3825)^2 - 1} \right)^2 \geq 0.859. \quad (\text{A.156})$$

Similarly, we also have

$$\sigma_{\min}^2(K_1^\top N_Z) \geq 0.859. \quad (\text{A.157})$$

3. We next derive useful expressions of A_{22} and \hat{A}_{22} . First we introduce the following quantities,

$$J_1^\top Z_{11,[1:\hat{r},1:\hat{r}]} K_1 \stackrel{(\text{A.145})}{=} J_1^\top B_{M,\hat{r}} K_1 + J_1^\top E_{M,\hat{r}} K_1 \triangleq B_{M1} + E_{M1}, \quad (\text{A.158})$$

$$J_2^\top Z_{11,[1:\hat{r},1:\hat{r}]} K_2 \stackrel{(\text{A.145})}{=} J_2^\top B_{M,\hat{r}} K_2 + J_2^\top E_{M,\hat{r}} K_2 \triangleq B_{M2} + E_{M2}, \quad (\text{A.159})$$

$$Z_{21,[:,1:\hat{r}]} K_1 \stackrel{(\text{A.146})}{=} B_{L,\hat{r}} K_1 + E_{L,\hat{r}} K_1 \triangleq B_{L1} + E_{L1}, \quad (\text{A.160})$$

$$Z_{21,[:,1:\hat{r}]} K_2 \stackrel{(\text{A.146})}{=} B_{L,\hat{r}} K_2 + E_{L,\hat{r}} K_2 \triangleq B_{L2} + E_{L2}, \quad (\text{A.161})$$

$$J_1^\top Z_{12,[1:\hat{r},:]} \stackrel{(\text{A.147})}{=} J_1^\top B_{R,\hat{r}} + J_1^\top E_{R,\hat{r}} \triangleq B_{R1} + E_{R1}, \quad (\text{A.162})$$

$$J_2^\top Z_{11,[1:\hat{r},:]} \stackrel{(\text{A.147})}{=} J_2^\top B_{R,\hat{r}} + J_2^\top E_{R,\hat{r}} \triangleq B_{R2} + E_{R2}. \quad (\text{A.163})$$

Since

$$\begin{aligned} & B_{L1} B_{M1}^{-1} B_{R1} = B_{L,\hat{r}} K_1 (J_1^\top B_{M,\hat{r}} K_1)^{-1} J_1^\top B_{R,\hat{r}} \\ & = U_{21} \Sigma_1 V_{11}^\top V_{[:,1:\hat{r}]}^{(1)} K_1 \left(J_1^\top U_{[:,1:\hat{r}]}^{(2)\top} U_{11} \Sigma_1 V_{11}^\top V_{[:,1:\hat{r}]}^{(1)} K_1 \right)^{-1} J_1^\top U_{[:,1:\hat{r}]}^{(2)\top} U_{11} \Sigma_1 V_{21}^\top \quad (\text{A.164}) \\ & = U_{21} \Sigma_1 V_{21}^\top, \end{aligned}$$

we can characterize A_{22} , \hat{A}_{22} by these new notations as

$$A_{22} = U_{21} \Sigma_1 V_{21}^\top + U_{22} \Sigma_2 V_{22}^\top \stackrel{(\text{A.164})}{=} B_{L1} B_{M1}^{-1} B_{R1} + U_{22} \Sigma_2 V_{22}^\top, \quad (\text{A.165})$$

$$\hat{A}_{22} = Z_{21,[:,1:\hat{r}]} Z_{11,[1:\hat{r},1:\hat{r}]}^{-1} Z_{12,[1:\hat{r},:]} \quad (\text{A.166})$$

$$\begin{aligned} &\stackrel{(\text{A.153})}{=} Z_{21,[:,1:\hat{r}]} K (J^\top Z_{11,[1:\hat{r},1:\hat{r}]} K)^{-1} J^\top Z_{12,[1:\hat{r},:]} \\ &= (Z_{21,[1:\hat{r}]} K_1 + Z_{21,[1:\hat{r}]} K_2) (J_1^\top Z_{11,[1:\hat{r},1:\hat{r}]} K_1 + J_2^\top Z_{11,[1:\hat{r},1:\hat{r}]} K_2)^{-1} \end{aligned} \quad (\text{A.167})$$

$$\begin{aligned} &\cdot (J_1^\top Z_{12,[1:\hat{r}]} + J_2^\top Z_{12,[1:\hat{r}]}) \\ &\stackrel{(\text{A.158})-(\text{A.163})}{=} \sum_{k=1}^2 (B_{Lk} + E_{Lk})(B_{Mk} + E_{Mk})^{-1}(B_{Rk} + E_{Rk}) \end{aligned} \quad (\text{A.168})$$

4. We now establish a number of bounds for the terms on the right hand side of (A.158)-(A.163).

Lemma A.3.8. *Based on the assumptions above, we have*

$$\sigma_{\min}(B_{M1}) \geq 3.43\sigma_{r+1}(A); \quad (\text{A.169})$$

$$\|B_{L1}B_{M1}^{-1}\| \leq \frac{\sqrt{3825/3824}}{\sqrt{0.859}\sigma_{\min}(U_{11})}, \quad \|B_{M1}^{-1}B_{R1}\| \leq \frac{\sqrt{3825/3824}}{\sqrt{0.859}\sigma_{\min}(V_{11})}, \quad (\text{A.170})$$

$$\|E_{Mt}\|_q \leq \|A_{-\max(r)}\|_q, \quad \|E_{Lt}\|_q \leq \|A_{-\max(r)}\|_q, \quad \|E_{Rt}\|_q \leq \|A_{-\max(r)}\|_q, \quad t = 1, 2, \quad (\text{A.171})$$

$$\|(B_{L2} + E_{L2})(B_{M2} + E_{M2})^{-1}\| \leq T_R + \frac{1}{1 - 1/3.43} \left(\frac{\sqrt{3825/3824}}{\sqrt{0.859}\sigma_{\min}(U_{11})} + \frac{1}{3.43} \right), \quad (\text{A.172})$$

$$\|B_{R2}\|_q \leq \frac{2\sqrt{3825/3824}}{\sqrt{0.859}\sigma_{\min}(V_{11})} \|A_{-\max(r)}\|_q. \quad (\text{A.173})$$

The proof of Lemma A.3.8 is given in the Supplement.

5. We finally give the upper bound of $\|\hat{A}_{22} - A_{22}\|_q$. By (A.165) and (A.168), we

can split the loss as,

$$\begin{aligned}\hat{A}_{22} - A_{22} &= ((B_{L1} + E_{L1})(B_{M1} + E_{M1})^{-1}(B_{R1} + E_{R1}) - B_{L1}B_{M1}^{-1}B_{R1}) \\ &\quad + (B_{L2} + E_{L2})(B_{M2} + E_{M2})^{-1}(B_{R2} + E_{R2}) - U_{22}\Sigma_2V_{22}^\top.\end{aligned}\tag{A.174}$$

We will analyze them separately. First, $\|U_{22}\Sigma_2V_{22}^\top\|_q \leq \|A_{-\max(r)}\|_q$; second,

$$\begin{aligned}&\|(B_{L2} + E_{L2})(B_{M2} + E_{M2})^{-1}(B_{R2} + E_{R2})\|_q \\ &\leq \|(B_{L2} + E_{L2})(B_{M2} + E_{M2})^{-1}\| \cdot (\|B_{R2}\|_q + \|E_{M2}\|_q) \\ &\stackrel{(A.172)(A.173)}{\leq} \left(T_R + \frac{3.43}{2.43} \left(\frac{\sqrt{3825/3824}}{\sqrt{0.859}\sigma_{\min}(U_{11})} + \frac{1}{3.43} \right) \right) \\ &\quad \cdot \left(\frac{2\sqrt{3825/3824}}{\sqrt{0.859}\sigma_{\min}(V_{11})} + 1 \right) \|A_{-\max(r)}\|_q \\ &\leq \left(T_R + \frac{1.524}{\sigma_{\min}(U_{11})} + 0.412 \right) \left(\frac{2.16}{\sigma_{\min}(V_{11})} + 1 \right) \|A_{-\max(r)}\|_q.\end{aligned}\tag{A.175}$$

The analysis of $((B_{L1} + E_{L1})(B_{M1} + E_{M1})^{-1}(B_{R1} + E_{R1}) - B_{L1}B_{M1}^{-1}B_{R1})$ is sim-

ilar to the proof of Theorem 3.3.1. We have

$$\begin{aligned}
& \left\| (B_{L1} + E_{L1})(B_{M1} + E_{M1})^{-1}(B_{R1} + E_{R1}) - B_{L1}B_{M1}^{-1}B_{R1} \right\|_q \\
& \leq \left\| B_{L1}(B_{M1}^{-1}E_{M1} \sum_{i=0}^{\infty} (-B_{M1}^{-1}E_{M1})^i B_{M1}^{-1})B_{R1} \right\|_q \\
& \quad + \left\| E_{L1} \left(\sum_{i=0}^{\infty} (-B_{M1}^{-1}E_{M1})^i B_{M1}^{-1} \right) B_{R1} \right\|_q \\
& \quad + \left\| B_{L1} \left(B_{M1}^{-1} \sum_{i=0}^{\infty} (-E_{M1}B_{M1}^{-1})^i \right) E_{R1} \right\|_q \\
& \quad + \left\| E_{L1} \left(B_{M1}^{-1} \sum_{i=0}^{\infty} (-E_{M1}B_{M1}^{-1})^i \right) E_{R1} \right\|_q \\
& \leq \|B_{L1}B_{M1}^{-1}\| \|E_{M1}\|_q \sum_{i=0}^{\infty} \|E_{M1}\|^i \|B_{M1}^{-1}\|^i \|B_{M1}^{-1}B_{R1}\| \\
& \quad + \|E_{L1}\|_q \sum_{i=0}^{\infty} \|B_{M1}^{-1}\|^i \|E_{M1}\|^i \|B_{M1}^{-1}B_{R1}\| \\
& \quad + \|B_{L1}B_{M1}^{-1}\| \sum_{i=0}^{\infty} \|E_{M1}\|^i \|B_{M1}^{-1}\|^i \|E_{R1}\|_q \\
& \quad + \|E_{L1}\| \sum_{i=0}^{\infty} \|B_{M1}^{-1}\|^{i+1} \|E_{M1}\|^i \|E_{R1}\|_q \\
& \stackrel{(A.171)}{\leq} \frac{\|\Sigma_2\|_q}{1 - \sigma_{r+1}(A)\|B_{M1}^{-1}\|} \\
& \quad (\|B_{L1}B_{M1}^{-1}\| \|B_{M1}^{-1}B_{R1}\| + \|B_{M1}^{-1}B_{R1}\| + \|B_{L1}B_{M1}^{-1}\| + \|B_{M1}^{-1}\| \sigma_{r+1}(A)) \\
& \stackrel{(A.170)(A.169)}{\leq} \left(\frac{1.65}{\sigma_{\min}(U_{11})\sigma_{\min}(V_{11})} + \frac{1.53}{\sigma_{\min}(V_{11})} + \frac{1.53}{\sigma_{\min}(V_{11})} + 0.42 \right) \|A_{-\max(r)}\|_q.
\end{aligned} \tag{A.176}$$

From (A.175), (A.176), (A.174), and the fact that $\sigma_{\min}(U_{11}) \leq 1$ and $T_R \geq$

$$\frac{1.36}{\sigma_{\min}(U_{11})} + 0.35,$$

$$\begin{aligned}
& \|\hat{A}_{22} - A_{22}\|_q \\
& \leq \left(2.16T_R + \left(\frac{4.95}{\sigma_{\min}(U_{11})} + 2.42 \right) \right) \left(\frac{1}{\sigma_{\min}(V_{11})} + 1 \right) \|A_{-\max(r)}\|_q \\
& \leq \left(2.16T_R + 4.31 \left(\frac{1.36}{\sigma_{\min}(U_{11})} + 0.35 \right) \right) \left(\frac{1}{\sigma_{\min}(V_{11})} + 1 \right) \|A_{-\max(r)}\|_q \\
& \leq 6.5T_R \left(\frac{1}{\sigma_{\min}(V_{11})} + 1 \right) \|A_{-\max(r)}\|_q.
\end{aligned} \tag{A.177}$$

This concludes the proof. \square

Proof of Lemma A.3.7.

In order to prove this lemma, we just need to prove that the for-loop in Algorithm 2 will break for some $s \geq r$. This can be shown by proving the break condition

$$\|D_{R,s}\| = \|Z_{21,[1:s]}Z_{11,[1:s,1:s]}^{-1}\| \leq T_R, \tag{A.178}$$

hold for $s = r$.

We adopt the definitions in (A.134), (A.135), (A.136), then we have

$$\begin{aligned}
Z_{11,[1:r,1:r]} &= U_{[:,1:r]}^{(2)\top} A_{11} V_{[:,1:r]}^{(1)} = \hat{M}^\top A_{11} \hat{N} \\
&= \hat{M}^\top U_{11} \Sigma_1 V_{11}^\top \hat{N} + \hat{M}^\top U_{12} \Sigma_2 V_{12}^\top \hat{N} \\
&= B_M + E_M,
\end{aligned}$$

$$Z_{21,[1:r]} = A_{21} V_{[:,1:r]}^{(1)} = (U_{21} \Sigma_1 V_{11}^\top + U_{22} \Sigma_2 V_{12}^\top) \hat{N} = B_L + E_L.$$

Hence,

$$\begin{aligned}
\left\| Z_{21,[:,1:r]} Z_{11,[1:r,1:r]}^{-1} \right\| &= \|(B_L + E_L)(B_M + E_M)^{-1}\| \\
&\leq \left\| B_L B_M^{-1} \sum_{i=0}^{\infty} (-E_M B_M^{-1})^i \right\| + \left\| E_L B_M^{-1} \sum_{i=0}^{\infty} (-E_M B_M^{-1})^i \right\| \\
&\leq (\|B_L B_M^{-1}\| + \|E_L\| \|B_M^{-1}\|) \frac{1}{1 - \|E_M B_M^{-1}\|} \\
&\stackrel{(A.139), (A.170)}{\leq} \left(\frac{\sqrt{45/44}}{\sigma_{\min}(U_{11})} + \frac{45\sigma_{r+1}(A)}{44\sigma_r(A)\sigma_{\min}(U_{11})\sigma_{\min}(V_{11})} \right) \\
&\quad \cdot \frac{1}{1 - \frac{45\sigma_{r+1}(A)}{44\sigma_r(A)\sigma_{\min}(U_{11})\sigma_{\min}(V_{11})}} \\
&\leq \frac{1.36}{\sigma_{\min}(U_{11})} + 0.35 \leq T_R,
\end{aligned}$$

which finished the proof of the lemma. \square

Proof of Lemma A.3.8.

First, since $M_Z \in \mathbb{R}^{\hat{r} \times r}$ and $N_Z \in \mathbb{R}^{\hat{r} \times r}$ are an orthonormal basis of $B_{M,\hat{r}}$ and $B_{M,\hat{r}}^\top$, we have $P_{M_Z} = M_Z M_Z^\top$ and $P_{N_Z} = N_Z N_Z^\top$ and

$$\begin{aligned}
\sigma_{\min}(B_{M1}) &= \sigma_{\min}(J_1^\top B_{M,\hat{r}} K_1) = \sigma_{\min}(J_1^\top M_Z M_Z^\top B_{M,\hat{r}} N_Z N_Z^\top K_1) \\
&\geq \sigma_{\min}(J_1^\top M_Z) \sigma_{\min}(M_Z^\top B_{M,\hat{r}} N_Z) \sigma_{\min}(N_Z^\top K_1) \\
&\stackrel{(A.156)(A.157)}{\geq} 0.859 \sigma_r(B_{M,\hat{r}}) \stackrel{(A.151)}{\geq} \frac{0.859 \cdot 3824}{3825} \sigma_r(A) \sigma_{\min}(U_{11}) \sigma_{\min}(V_{11}) \stackrel{(3.12)}{\geq} 3.43 \sigma_{r+1}(A).
\end{aligned} \tag{A.179}$$

which gives (A.169).

$$\begin{aligned}
& \|B_{L1}B_{M1}^{-1}\| = \left\| B_{L,\hat{r}}K_1 (J_1^\top B_{M,\hat{r}}K_1)^{-1} \right\| \\
& = \left\| U_{21}\Sigma_1 V_{11}^\top V_{[:,1:\hat{r}]}^{(1)} K_1 \left(J_1^\top U_{[:,1:\hat{r}]}^{(2)\top} U_{11}\Sigma_1 V_{11}^\top V_{[:,1:\hat{r}]}^{(1)} K_1 \right)^{-1} \right\| = \left\| U_{21} \left(J_1^\top U_{[:,1:\hat{r}]}^{(2)\top} U_{11} \right)^{-1} \right\| \\
& \leq \frac{1}{\sigma_{\min}(J_1^\top U_{[:,1:\hat{r}]}^{(2)\top} U_{11})} = \frac{1}{\sigma_{\min}(J_1^\top P_{M_Z}(U_{[:,1:\hat{r}]}^{(2)\top} U_{11}))} = \frac{1}{\sigma_{\min}((J_1^\top M_Z)(M_Z^\top U_{[:,1:\hat{r}]}^{(2)\top} U_{11}))} \\
& \leq \frac{1}{\sigma_{\min}(J_1^\top M_Z)} \cdot \frac{1}{\sigma_{\min}(U_{[:,1:\hat{r}]}^{(2)\top} U_{11})} \stackrel{(A.149)(A.156)}{\leq} \frac{\sqrt{3825/3824}}{\sqrt{0.859}\sigma_{\min}(U_{11})},
\end{aligned} \tag{A.180}$$

which gives the first part of (A.170). Here we used the fact that $\Sigma_1 V_{11}^\top V_{[:,1:\hat{r}]}^{(1)} K_1$ is a square matrix; M_Z is the orthonormal basis of the column space of $Z_{11,[1:\hat{r},1:\hat{r}]} = U_{[:,1:\hat{r}]}^{(2)\top} U_{11}\Sigma_1 V_{11}^\top V_{[:,1:\hat{r}]}^{(1)}$. Similarly we have the later part of (A.170),

$$\|B_{M1}^{-1}B_{R1}\| \leq \frac{\sqrt{3825/3824}}{\sqrt{0.859}\sigma_{\min}(V_{11})}. \tag{A.181}$$

Based on the definitions, we have the bound for all “ E ” terms in (A.158)-(A.163), i.e. (A.171). Now we move on to (A.172). By the SVD of $Z_{11,[1:\hat{r},1:\hat{r}]}$ (A.153) and the partition (A.154), we know

$$\begin{aligned}
& ([J_1 \ J_2]^\top Z_{11,[1:\hat{r},1:\hat{r}]} [K_1 \ K_2])^{-1} = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix}^{-1} \\
& = \begin{bmatrix} (J_1^\top Z_{11,[1:\hat{r},1:\hat{r}]} K_1)^{-1} & 0 \\ 0 & (J_2^\top Z_{11,[1:\hat{r},1:\hat{r}]} K_2)^{-1} \end{bmatrix}.
\end{aligned}$$

Hence, we have

$$\begin{aligned}
& \left\| (B_{L2} + E_{L2})(B_{M2} + E_{M2})^{-1} \right\| = \left\| Z_{21,[:,1:\hat{r}]} K_2 (J_2^\top Z_{11,[:,1:\hat{r}]} K_2)^{-1} \right\| \\
& = \left\| Z_{21,[:,1:\hat{r}]} [K_1 \ K_2] \left([J_1 \ J_2]^\top Z_{11,[:,1:\hat{r}]} [K_1 \ K_2] \right)^{-1} \right. \\
& \quad \left. - Z_{21,[:,1:\hat{r}]} K_1 (J_1^\top Z_{11,[:,1:\hat{r}]} K_1)^{-1} \right\| \\
& \leq \left\| Z_{21,[:,1:\hat{r}]} (Z_{11,[:,1:\hat{r}]})^{-1} \right\| + \left\| (B_{L1} + E_{L1})(B_{M1} + E_{M1})^{-1} \right\| \\
& \leq T_R + \left\| B_{L1} \cdot B_{M1}^{-1} \sum_{i=0}^{\infty} (-E_{M1} B_{M1}^{-1})^i \right\| + \left\| E_{L1} \cdot B_{M1}^{-1} \sum_{i=0}^{\infty} (-E_{M1} B_{M1}^{-1})^i \right\| \\
& \leq T_R + \left(\|B_{L1} B_{M1}^{-1}\| + \|E_{L1}\| \|B_{M1}^{-1}\| \right) \frac{1}{1 - \|E_{M1}\| \|B_{M1}^{-1}\|} \\
& \stackrel{(A.169)(A.170)(A.171)}{\leq} T_R + \left(\frac{\sqrt{3825/3824}}{\sqrt{0.859} \sigma_{\min}(U_{11})} + \frac{1}{3.43} \right) \cdot \frac{1}{1 - 1/3.43},
\end{aligned} \tag{A.182}$$

which proves (A.172). Since $Z_{11,[:,1:\hat{r}]} = B_{M,\hat{r}} + E_{M,\hat{r}}$ and by definition, $\text{rank}(B_{M,\hat{r}}) \leq r$, by Lemma A.3.1, we know

$$\sigma_{r+i}(Z_{11,[:,1:\hat{r}]}) \leq \sigma_i(E_{M,\hat{r}}), \quad \forall i \geq 1. \tag{A.183}$$

Then

$$\begin{aligned}
\|B_{M2}\|_q & \leq \|B_{M2} + E_{M2}\|_q + \|E_{M2}\|_q \leq \|J_2^\top Z_{11,[:,1:\hat{r}]} K_2\|_q + \|E_{M2}\|_q \\
& = \sqrt[q]{\sum_{i=r+1}^{\hat{r}} \sigma_i^q(Z_{11,[:,1:\hat{r}]})} + \|E_{M2}\|_q \leq \sqrt[q]{\sum_{i=1}^{\hat{r}-r} \sigma_i^q(E_{M,\hat{r}})} + \|E_{M2}\|_q \tag{A.184} \\
& \leq \|E_{M,\hat{r}}\|_q + \|E_{M2}\|_q \stackrel{(A.171)}{\leq} 2\|A_{-\max(r)}\|_q.
\end{aligned}$$

Same to the process of (A.180), we know

$$\frac{1}{\sigma_{\min}(V_{11}^\top V_{[:,1:\hat{r}]}^{(1)} K_1)} \leq \frac{\sqrt{3825/3824}}{\sqrt{0.859} \sigma_{\min}(V_{11})}. \tag{A.185}$$

Also, $\|V_{21}^\top\| \leq 1$. Hence,

$$\begin{aligned}
\|B_{R2}\|_q &\stackrel{(A.163)}{=} \|J_2^\top B_{R,\hat{r}}\|_q = \|J_2^\top U_{[:,1:\hat{r}]}^{(2)\top} U_{11} \Sigma_1 V_{21}^\top\|_q \\
&= \|J_2^\top U_{[:,1:\hat{r}]}^{(2)\top} U_{11} \Sigma_1 (V_{11}^\top V_{[:,1:\hat{r}]}^{(1)} K_1) (V_{11}^\top V_{[:,1:\hat{r}]}^{(1)} K_1)^{-1} V_{21}^\top\|_q \\
&\leq \|B_{M2}\|_q \cdot \|(V_{11}^\top V_{[:,1:\hat{r}]}^{(1)} K_1)^{-1}\| \cdot \|V_{21}^\top\| \\
&\stackrel{(A.184)(A.185)}{\leq} \frac{2\sqrt{3825/3824}}{\sqrt{0.859}\sigma_{\min}(V_{11})} \|A_{-\max(r)}\|_q.
\end{aligned} \tag{A.186}$$

which proves (A.173). \square

Proof of Theorem 3.3.3.

The idea of proof is to construct two matrices $A^{(1)}, A^{(2)}$ both in $\mathcal{F}_c(M_1, M_2)$ such that they have the identical first m_1 rows and m_2 columns, but differ much in the remaining block. Suppose $a, b, c > 0$ are fixed numbers, ε is a small real number. We first consider the following 2-by-2 matrix

$$B(\varepsilon) = \begin{bmatrix} a & c \\ b & \frac{bc}{a} + \varepsilon \end{bmatrix}. \tag{A.187}$$

Suppose the larger and smaller singular value of $B(\varepsilon)$ are $\lambda_{\max}(\varepsilon)$ and $\lambda_{\min}(\varepsilon)$, then we have

$$\lambda_{\max}(\varepsilon) \rightarrow \|B(0)\| = \frac{\sqrt{(a^2 + b^2)(a^2 + c^2)}}{a} \tag{A.188}$$

as $\varepsilon \rightarrow 0$; since $\lambda_{\max}(\varepsilon) \cdot \lambda_{\min}(\varepsilon) = |\det(B)| = a|\varepsilon|$, we also have

$$\lambda_{\min}(\varepsilon)/|\varepsilon| \rightarrow \frac{a^2}{\sqrt{(a^2 + b^2)(a^2 + c^2)}} \tag{A.189}$$

as $\varepsilon \rightarrow 0$. If $B(\varepsilon)$ defined in (A.187) has SVD

$$B(\varepsilon) = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{21} \end{bmatrix} \cdot \begin{bmatrix} \lambda_{\max}(\varepsilon) & 0 \\ 0 & \lambda_{\min}(\varepsilon) \end{bmatrix} \cdot \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{21} \end{bmatrix}^\top \quad (\text{A.190})$$

then we also have

$$u_{11} \rightarrow \frac{a}{\sqrt{a^2 + b^2}}, \quad u_{21} \rightarrow \frac{b}{\sqrt{a^2 + b^2}}, \quad v_{11} \rightarrow \frac{a}{\sqrt{a^2 + c^2}}, \quad v_{21} \rightarrow \frac{c}{\sqrt{a^2 + c^2}}. \quad (\text{A.191})$$

as $\varepsilon \rightarrow 0$.

Now we set $a = 1$, $b = \sqrt{1 - M_1^2}/M_1 - \eta$, $c = \sqrt{1 - M_2^2}/M_2 - \eta$, $d = bc/a$, where η is some small positive number to be specify later. We construct $A_{11}, A_{12}, A_{21}, A_{22}^{(1)}$ and $A_{22}^{(2)}$ such that,

$$A_{11} = \begin{bmatrix} aI_r & 0 \\ 0 & 0 \end{bmatrix}_{m_1 \times m_2}, \quad A_{12} = \begin{bmatrix} cI_r & 0 \\ 0 & 0 \end{bmatrix}_{m_1 \times (p_2 - m_2)}, \quad A_{21} = \begin{bmatrix} bI_r & 0 \\ 0 & 0 \end{bmatrix}_{(p_1 - m_1) \times m_2}; \quad (\text{A.192})$$

$$A_{22}^{(1)} = \begin{bmatrix} (d + \varepsilon)I_r & 0 \\ 0 & 0 \end{bmatrix}_{(p_1 - m_1) \times (p_2 - m_2)}, \quad A_{22}^{(2)} = \begin{bmatrix} (d - \varepsilon)I_r & 0 \\ 0 & 0 \end{bmatrix}_{(p_1 - m_1) \times (p_2 - m_2)}. \quad (\text{A.193})$$

Here we use I_r to note the identity matrix of dimension r . Then we construct $A^{(1)}$ and $A^{(2)}$ as

$$A^{(1)} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}^{(1)} \end{bmatrix}, \quad A^{(2)} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}^{(2)} \end{bmatrix}, \quad (\text{A.194})$$

where $A^{(1)}$ and $A^{(2)}$ are with identical first m_1 rows and m_2 columns. Since the SVD

of $B(\varepsilon)$ is given as (A.190), the SVD of $A^{(1)}$ can be written as

$$A^{(1)} = \begin{bmatrix} U_{11}^{(1)} & U_{12}^{(1)} \\ U_{21}^{(1)} & U_{22}^{(1)} \end{bmatrix} \cdot \begin{bmatrix} \Sigma_1^{(1)} & 0 \\ 0 & \Sigma_2^{(1)} \end{bmatrix} \cdot \begin{bmatrix} V_{11}^{(1)} & V_{12}^{(1)} \\ V_{21}^{(1)} & V_{22}^{(1)} \end{bmatrix}^\top,$$

where

$$\begin{aligned} U_{11} &= \begin{bmatrix} u_{11}I_r \\ 0 \end{bmatrix}_{m_1 \times r}, & U_{12} &= \begin{bmatrix} u_{12}I_r \\ 0 \end{bmatrix}_{m_1 \times r}, \\ U_{21} &= \begin{bmatrix} u_{21}I_r \\ 0 \end{bmatrix}_{(p_1-m_1) \times r}, & U_{22} &= \begin{bmatrix} u_{22}I_r \\ 0 \end{bmatrix}_{(p_1-m_1) \times r}; \\ V_{11} &= \begin{bmatrix} v_{11}I_r \\ 0 \end{bmatrix}_{m_2 \times r}, & V_{12} &= \begin{bmatrix} v_{12}I_r \\ 0 \end{bmatrix}_{m_2 \times r}, \\ V_{21} &= \begin{bmatrix} v_{21}I_r \\ 0 \end{bmatrix}_{(p_2-m_2) \times r}, & V_{22} &= \begin{bmatrix} v_{22}I_r \\ 0 \end{bmatrix}_{(p_2-m_2) \times r}; \\ \Sigma_1 &= \lambda_{\max}(\varepsilon)I_r, & \Sigma_2 &= \lambda_{\min}(\varepsilon)I_r. \end{aligned}$$

Hence,

$$\sigma_{\min}(U_{11}) = u_{11} = \frac{a}{\sqrt{a^2 + b^2}} \rightarrow \frac{1}{1 + \left(\frac{\sqrt{1-M_1^2}}{M_1} - \eta\right)^2} > M_1, \quad \text{as } \varepsilon \rightarrow 0$$

$$\sigma_{\min}(V_{11}) = v_{11} = \frac{a}{\sqrt{a^2 + c^2}} \rightarrow \frac{1}{1 + \left(\frac{\sqrt{1-M_2^2}}{M_2} - \eta\right)^2} > M_2, \quad \text{as } \varepsilon \rightarrow 0.$$

Also, $\|\Sigma_2^{(1)}\| \rightarrow 0$ as $\varepsilon \rightarrow 0$. So we have $A^{(1)} \in \mathcal{F}_r(M_1, M_2)$ when ε is small enough.

Similarly $A^{(2)} \in \mathcal{F}_r(M_1, M_2)$ when ε is small enough. Now we also have $\|A_{-\max(r)}^{(1)}\|_q = (q\lambda_{\min}(\varepsilon)^q)^{1/q} = q^{1/q}\lambda_{\min}(\varepsilon)$, $\|A_{-\max(r)}^{(2)}\|_q = (q\lambda_{\min}(-\varepsilon)^q)^{1/q} = q^{1/q}\lambda_{\min}(-\varepsilon)$. $\|A_{22}^{(1)} -$

$$A_{22}^{(2)}\|_q = (q(2|\varepsilon|)^q)^{1/q} = 2|\varepsilon|q^{1/q}.$$

Finally for any estimate \hat{A}_{22} , we must have

$$\begin{aligned} & \max \left\{ \frac{\|\hat{A}_{22} - A_{22}^{(1)}\|_q}{\|A_{-\max(r)}^{(1)}\|_q}, \frac{\|\hat{A}_{22} - A_{22}^{(2)}\|_q}{\|A_{-\max(r)}^{(2)}\|_q} \right\} \geq \frac{\frac{1}{2} \left\| \left(\hat{A}_{22} - A_{22}^{(1)} \right) - \left(\hat{A}_{22} - A_{22}^{(2)} \right) \right\|_q}{\min \left\{ \|A_{-\max(r)}^{(1)}\|_q, \|A_{-\max(r)}^{(2)}\|_q \right\}} \\ & \geq \frac{2|\varepsilon|}{2 \min \left\{ \lambda_{\min}(\varepsilon), \lambda_{\min}(-\varepsilon) \right\}} \stackrel{\text{(A.189)}}{\rightarrow} \frac{\sqrt{(a^2 + b^2)(a^2 + c^2)}}{a^2} \\ & = \sqrt{\left(1 + \left(\frac{\sqrt{1 - M_1^2}}{M_1} - \eta \right)^2 \right) \left(1 + \left(\frac{\sqrt{1 - M_2^2}}{M_2} - \eta \right)^2 \right)} \end{aligned} \tag{A.195}$$

as $\varepsilon \rightarrow 0$. Since $A^{(1)}, A^{(2)} \in \mathcal{F}_r(M_1, M_2)$ and are with identical first m_1 rows and m_2 columns, we must have

$$\inf_{\hat{A}_{22}} \sup_{A \in \mathcal{F}_r(M_1, M_2)} \frac{\|\hat{A}_{22} - A_{22}\|_q}{\|A_{-\max(r)}\|_q} \geq \sqrt{\left(1 + \left(\frac{\sqrt{1 - M_1^2}}{M_1} - \eta \right)^2 \right) \left(1 + \left(\frac{\sqrt{1 - M_2^2}}{M_2} - \eta \right)^2 \right)}.$$

Let $\eta \rightarrow 0$, since $M_1, M_2 < 1$, we have

$$\inf_{\hat{A}_{22}} \sup_{A \in \mathcal{F}_r(M_1, M_2)} \frac{\|\hat{A}_{22} - A_{22}\|_q}{\|A_{-\max(r)}\|_q} \geq \frac{1}{M_1 M_2} \geq \frac{1}{4} \left(\frac{1}{M_1} + 1 \right) \left(\frac{1}{M_2} + 1 \right), \tag{A.196}$$

which finished the proof of theorem. \square

Proof of Corollary 3.3.1.

We first prove the second part of the corollary. We set $\alpha = (136/165)^2$. Since $U_{[:,1:r]} \in \mathbb{R}^{p_1 \times r}$ is with orthonormal columns, by Lemma A.3.5 and

$$m_1 \geq 12.5 W_r^{(1)} r (\log r + c) \geq \frac{4}{(1 - \alpha)^2} \cdot W_r^{(1)} r (\log r + c),$$

we have

$$\sigma_{\min}(U_{11}) = \sigma_{\min}(U_{[\Omega_1, 1:r]}) \geq \sqrt{\frac{\alpha m_1}{p_1}} \quad (\text{A.197})$$

with probability at least $1 - 2 \exp(-c)$. When (A.197) holds, by the condition, we know

$$\begin{aligned} \sigma_{r+1}(A) &\leq \sigma_r(A) \sigma_{\min}(V_{11}) \frac{1}{5} \sqrt{\frac{m_1}{p_1}} \leq \sigma_r(A) \sigma_{\min}(V_{11}) \frac{1}{5\sqrt{\alpha}} \cdot \sigma_{\min}(U_{11}) \\ &\leq \frac{1}{4} \sigma_r(A) \sigma_{\min}(V_{11}) \sigma_{\min}(U_{11}). \end{aligned}$$

When $T_R \geq 2\sqrt{p_1/m_1}$, we have

$$\frac{1.36}{\sigma_{\min}(U_{11})} + 0.35 \leq 1.36 \sqrt{\frac{p_1}{\alpha m_1}} + 0.35 \leq 2\sqrt{\frac{p_1}{m_1}} \leq T_R$$

Hence we can apply Theorem 3.3.2, for $1 \leq q \leq \infty$ we must have

$$\left\| \hat{A}_{22} - A_{22} \right\|_q \leq 6.5 T_R \left\| A_{-\max(r)} \right\|_q \left(\frac{1}{\sigma_{\min}(V_{11})} + 1 \right), \quad (\text{A.198})$$

which finishes the proof of the second part of Corollary 3.3.1. Besides, the proof for the third part is the same as the second part after we take the transpose of the matrix.

For the first part, the proof is also similar. Again we set $\alpha = (136/165)^2$. Then we have

$$m_1 \geq \frac{4}{(1-\alpha)^2} W_r^{(1)} r (\log r + c), \quad m_2 \geq \frac{4}{(1-\alpha)^2} W_r^{(2)} r (\log r + c),$$

so

$$\sigma_{\min}(U_{11}) = \sigma_{\min}(U_{[\Omega_1, 1:r]}) \geq \sqrt{\frac{\alpha m_1}{p_1}}, \quad \sigma_{\min}(V_{11}) = \sigma_{\min}(V_{[\Omega_2, 1:r]}) \geq \sqrt{\frac{\alpha m_2}{p_2}} \quad (\text{A.199})$$

with probability at least $1 - 4 \exp(-c)$. When (A.199) holds, we have

$$\sigma_{r+1}(A) \leq \sigma_r(A) \frac{1}{6} \sqrt{\frac{m_1 m_2}{p_1 p_2}} \leq \sigma_r(A) \frac{1}{6\alpha} \sigma_{\min}(U_{11}) \sigma_{\min}(V_{11}) \leq \frac{1}{4} \sigma_r(A) \sigma_{\min}(V_{11}) \sigma_{\min}(U_{11}).$$

When $T_R = 2\sqrt{p_1/m_1}$ or $T_C = 2\sqrt{p_2/m_2}$, similarly to the first part we have

$$\frac{1.36}{\sigma_{\min}(U_{11})} + 0.35 \leq T_R, \quad \text{or} \quad \frac{1.36}{\sigma_{\min}(V_{11})} + 0.35 \leq T_C.$$

Hence we can apply Theorem 3.3.2 and get

$$\begin{aligned} \left\| \hat{A}_{22} - A_{22} \right\|_q &\leq 6.5 T_R \|A_{-\max(r)}\|_q \left(\frac{1}{\sigma_{\min}(V_{11})} + 1 \right) \\ &\leq 6.5 \cdot 2 \sqrt{\frac{p_1}{m_1}} \cdot \left(\sqrt{\frac{p_2}{\alpha m_2}} + 1 \right) \|A_{-\max(r)}\|_q \\ &\leq 29 \|A_{-\max(r)}\|_q \sqrt{\frac{p_1 p_2}{m_1 m_2}}. \end{aligned}$$

□

Proof of Corollary 3.3.2.

Suppose $0 < \alpha_1 < 1$, since $U_{[:,1:r]} \in \mathbb{R}$ is with random orthonormal columns of Haar measure, we can apply Lemma A.3.6 and find some $c > 0$ and $\delta > 0$ such that when $p_1 \geq m_1 \geq cr$,

$$\sigma_{\min}(U_{11}) = \sigma_{\min}(U_{[1:m_1, 1:r]}) \geq \frac{136}{165} \sqrt{\frac{m_1}{p_1}} \quad (\text{A.200})$$

with probability at least $1 - \exp(-\delta m_1)$. When (A.200) happen, we have

$$\begin{aligned} \sigma_{r+1}(A) &\leq \sigma_r(A) \sigma_{\min}(V_{11}) \frac{1}{5} \sqrt{\frac{m_1}{p_1}} \leq \sigma_r(A) \sigma_{\min}(V_{11}) \sigma_{\min}(U_{11}), \\ \frac{1.36}{\sigma_{\min}(U_{11})} + 0.35 &\leq 1.36 \cdot \frac{165}{136} \sqrt{\frac{p_1}{m_1}} + 0.35 \leq 2 \sqrt{\frac{p_1}{m_1}}. \end{aligned}$$

Hence we can apply Theorem 3.3.2, for $1 \leq q \leq \infty$, we have

$$\left\| \hat{A}_{22} - A_{22} \right\|_q \leq 6.5 T_R \left\| A_{-\max(r)} \right\|_q \left(\frac{1}{\sigma_{\min}(V_{11})} + 1 \right), \quad (\text{A.201})$$

which finishes the proof of the corollary. \square

Description of Cross-Validation

In this section, we describe the cross-validation used in penalized nuclear norm minimization (3.4) in the numerical comparison in Sections 3.4 and 3.5.

First, we construct a grid T of non-negative numbers based on a pre-selected positive integer N . Denote

$$t_{\max}^{PN} = \left\| \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & 0 \end{bmatrix} \right\|,$$

i.e. the largest singular value of the observed blocks. For penalized nuclear norm minimization, we let $T = \{t_{\max}^{PN}, t_{\max}^{PN} \cdot 10^{-3(1/N)}, \dots, t_{\max}^{PN} \cdot 10^{-3(N/N)}\}$.

Next, for a given positive integer K , we randomly divide the integer set $\{1, \dots, m_1\}$ into two groups of size $m^{(1)} \approx \frac{(K-1)n}{K}$, $m^{(2)} \approx \frac{n}{K}$ for H times. For $h = 1, \dots, H$, we denote by J_1^h and $J_2^h \subseteq \{1, 2, \dots, m_1\}$ the index sets of the two groups for the h -th split. Then the penalized nuclear norm minimization estimator (3.4) is applied to the first group of data: $A_{11}, A_{21}, (A_{12})_{[J_1^h, :]}$, i.e. the data of the observation set $\Omega = \{(i, j) : 1 \leq j \leq m_2, \text{ or } i \in J_1^h, m_2 + 1 \leq j \leq p_2\}$, with each value of the tuning parameter $t \in T$ and denote the result by $\hat{A}_h^{PN}(t)$. Note that we did not use the observed block $A_{[J_2^h, (m_2+1):p_2]}$ in calculating $\hat{A}_h^{PN}(t)$. Instead, $A_{[J_2^h, (m_2+1):p_2]}$ is used to

evaluate the performance of the tuning parameter $t \in T$. Set

$$\hat{R}(t) = \frac{1}{H} \sum_{h=1}^H \left\| \left[\hat{A}_h^{PN}(t) \right]_{[J_2^h, (m_2+1):p_2]} - A_{[J_2^h, (m_2+1):p_2]} \right\|_F^2. \quad (\text{A.202})$$

Finally, the tuning parameter is chosen as

$$t_* = \arg \min_{t \in T} \hat{R}(t)$$

and the final estimator \hat{A}^{PN} is calculated using this choice of the tuning parameter t_* .

In all the numerical studies with penalized nuclear norm minimization in Sections 3.4 and 3.5, we use 5-cross-validation (i.e., $K = 5$), $N = 10$ to select the tuning parameter.

Bibliography

- Alquier, P., Butucea, C., Hebiri, M., and Meziani, K. (2013). Rank penalized estimation of a quantum system. *Phys. Rev. A*, 88:032113.
- Andersson, J. and Stromberg, J.-O. (2014). On the theorem of uniform recovery of random sampling matrices. *IEEE Trans. Inf. Theory*, 60(3):1700–1710.
- Andrews, H. C. and Patterson, C. L. I. (1976). Singular value decomposition (svd) image coding. *IEEE Trans. Commun.*, 24:425–432.
- Argyriou, A., Evgeniou, T., and Pontil, M. (2008). Convex multi-task feature learning. *Mach. Learn.*, 73(3):243–272.
- Baraniuk, R., Davenport, M., DeVore, R., and M.Wakin (2008). A simple proof of the restricted isometry property for random matrices. *Constr. Approx.*, 28:253–263.
- Baraniuk, R. and Steeghs, P. (2007). Compressive radar imaging. *Proc. IEEE Radar Conf.*, pages 128–133.
- Basri, R. and Jacobs, D. W. (2003). Lambertian reflectance and linear sub-spaces. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25:218–233.
- Berchuck, A., Iversen, E. S., Lancaster, J. M., Pittman, J., Luo, J., Lee, P., Murphy, S., Dressman, H. K., Febbo, P. G., West, M., et al. (2005). Patterns of gene expression that characterize long-term survival in advanced stage serous ovarian cancers. *Clin. Cancer Res.*, 11(10):3686–3696.
- Birnbaum, A., Johnstone, I., Nadler, B., and Paul, D. (2013). Minimax bounds for sparse pca with noisy high-dimensional data. *Ann. Statist.*, 41:1055–1084.
- Biswas, P., Lian, T.-C., Wang, T.-C., and Ye, Y. (2006). Semidefinite programming based algorithms for sensor network localization. *ACM T. Sensor Network*, 2(2):188–220.
- Bonome, T., Lee, J.-Y., Park, D.-C., Radonovich, M., Pise-Masison, C., Brady, J., Gardner, G. J., Hao, K., Wong, W. H., Barrett, J. C., et al. (2005). Expression profiling of serous low malignant potential, low-grade, and high-grade tumors of the ovary. *Cancer Res.*, 65(22):10602–10612.

- Browning, B. L. and Browning, S. R. (2009). A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am. J. Hum. Genet.*, 84(2):210–223.
- Cai, J.-F., Candès, E., and Shen, Z. (2010a). A singular value thresholding algorithm for matrix completion. *SIAM J. Optim.*, 20(4):1956–1982.
- Cai, T., Cai, T. T., and Zhang, A. (2015). Structured matrix completion with applications to genomic data integration. *J. Amer. Statist. Assoc.*, to appear.
- Cai, T. T., Ma, Z., and Wu, Y. (2013a). Optimal estimation and rank detection for sparse spiked covariance matrices. *Probab. Theory Related Fields*, pages 1–35.
- Cai, T. T., Ma, Z., and Wu, Y. (2013b). Sparse PCA: Optimal rates and adaptive estimation. *Ann. Statist.*, 41:3074–3110.
- Cai, T. T., Wang, L., and Xu, G. (2010b). New bounds for restricted isometry constants. *IEEE Trans. Inf. Theory*, 56:4388–4394.
- Cai, T. T., Wang, L., and Xu, G. (2010c). Shifting inequality and recovery of sparse signals. *IEEE Trans. Signal Process.*, 58:1300–1308.
- Cai, T. T., Wang, L., and Xu, G. (2010d). Stable recovery of sparse signals and an oracle inequality. *IEEE Trans. Inf. Theory*, 56:3516–3522.
- Cai, T. T., Xu, G., and Zhang, J. (2009). On recovery of sparse signal via l_1 minimization. *IEEE Trans. Inf. Theory*, 55:3388–3397.
- Cai, T. T. and Zhang, A. (2013a). Compressed sensing and affine rank minimization under restricted isometry. *IEEE Trans. Signal Process.*, 61:3279–3290.
- Cai, T. T. and Zhang, A. (2013b). Sharp RIP bound for sparse signal and low-rank matrix recovery. *Appl. and Comput. Harmon. Anal.*, 35:74–93.
- Cai, T. T. and Zhang, A. (2014a). Perturbation bound on unilateral singular vectors. *Technical report*.
- Cai, T. T. and Zhang, A. (2014b). Sparse representation of a polytope and recovery of sparse signals and low-rank matrices. *IEEE Trans. Inf. Theory*, 60:122–132.
- Cai, T. T. and Zhang, A. (2015). ROP: Matrix recovery via rank-one projections. *Ann. Statist.*, 43:102–138.
- Cai, T. T. and Zhou, W. (2013). Matrix completion via max-norm constrained optimization. *arXiv preprint arXiv:1303.0341*.
- Cancer Genome Atlas Research Network (2011). Integrated genomic analyses of ovarian carcinoma. *Nature*, 474(7353):609–615.
- Candès, E. J. (2008). The restricted isometry property and its implications for compressed sensing. *C. R. Acad. Sci. Serie I*, 346:589–592.
- Candès, E. J. and Plan, Y. (2011). Tight oracle for low-rank matrix recovery from a minimal number of random measurements. *IEEE Trans. Inf. Theory*, 57:2342–2359.

- Candès, E. J. and Recht, B. (2009). Exact matrix completion via convex optimization. *Found. Comput. Math.*, 9:717–772.
- Candès, E. J., Romberg, J., and Tao, T. (2006). Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math.*, 59:1207–1223.
- Candès, E. J., Strohmer, T., and Voroninski, V. (2013). Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. *Commun. Pur. Appl. Math.*, 66:1241–1274.
- Candès, E. J. and Tao, T. (2005). Decoding by linear programming. *IEEE Trans. Inf. Theory*, 51:4203–4215.
- Candès, E. J. and Tao, T. (2007). The dantzig selector: statistical estimation when p is much larger than n . *Ann. Statist.*, 35:2313–2351.
- Candès, E. J. and Tao, T. (2010). The power of convex relaxation: Near-optimal matrix completion. *IEEE Trans. Inform. Theory*, 56:2053 – 2080.
- Candès, E. J., Li, X., Ma, Y., and Wright, J. (2011). Robust principal component analysis. *J. ACM*, 58:1–37.
- Chen, P. and Suter, D. (2004). Recovering the missing components in a large noisy low-rank matrix: Application to sfm. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(8):1051–1063.
- Chen, Y., Chi, Y., and Goldsmith, A. (2013). Exact and stable covariance estimation from quadratic sampling via convex programming. *arxiv preprint arXiv:1310.0807*.
- Chi, E. C., Zhou, H., Chen, G. K., Del Vecchio, D. O., and Lange, K. (2013). Genotype imputation via matrix completion. *Genome Res.*, 23(3):509–518.
- Dasarathy, G., Shah, P., Bhaskar, B. N., and Nowak, R. (2012). Covariance sketching. *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on. IEEE*.
- Dasarathy, G., Shah, P., Bhaskar, B. N., and Nowak, R. (2013). Sketching sparse matrices. *arXiv preprint arXiv:1303.6544*.
- Davenport, M., Laska, J., Treichler, J., and Baraniuk, R. (2012). The pros and cons of compressive sensing for wideband signal acquisition: Noise folding versus dynamic range. *IEEE Trans. Signal Process.*, 60:4628–4642.
- Davies, M. and Gribonval, R. (2009). Restricted isometry constants where l_p sparse recovery can fail for $0 < p \leq 1$. *IEEE Trans. Inf. Theory*, 55.
- Denkert, C., Budczies, J., Darb-Esfahani, S., Györfy, B., Schouli, J., Könsgen, D., Zeillinger, R., Weichert, W., Noske, A., Buckendahl, A.-C., et al. (2009). A prognostic gene expression index in ovarian cancer validation across different independent data sets. *J. Pathol.*, 218(2):273–280.
- Donoho, D. L., Elad, M., and Temlyakov, V. N. (2006). Stable recovery of sparse overcomplete representations in the presence of noise. *IEEE Trans. Inf. Theory*, 52:6–18.
- Donoho, D. L. and Huo, X. (2001). Uncertainty principles and ideal atomic decomposition. *IEEE Trans. Inf. Theory*, 47:2845–2862.
- Donoho, D. L. and Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81:425–455.

- Dressman, H. K., Berchuck, A., Chan, G., Zhai, J., Bild, A., Sayer, R., Cragun, J., Clarke, J., Whitaker, R. S., Li, L., et al. (2007). An integrated genomic-based approach to individualized treatment of patients with advanced-stage ovarian cancer. *J. Clin. Oncol.*, 25(5):517–525.
- Dvijotham, K. and Fazel, M. (2010). A nullspace analysis of the nuclear norm heuristic for rank minimization. *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pages 3586–3589.
- Fan, J., Fan, Y., and Lv, J. (2008). High dimensional covariance matrix estimation using a factor model. *J. Econom.*, 147:186–197.
- Foygel, R., Salakhutdinov, R., Shamir, O., and Srebro, N. (2011). Learning with the weighted trace-norm under arbitrary sampling distributions. In *NIPS*, pages 2133–2141.
- Fuchs, J.-J. (2004). On sparse representations in arbitrary redundant bases. *IEEE Trans. Inf. Theory*, 50:1341–1344.
- Grant, M. and Boyd, S. (2008). *Graph implementations for nonsmooth convex programs, Recent Advances in Learning and Control (a tribute to M. Vidyasagar). (V. Blondel, et al eds.), Lecture Notes in Control and Information Sciences*. Springer.
- Grant, M. and Boyd, S. (2012). Cvx: Matlab software for disciplined convex programming, version 2.0 beta, <http://cvxr.com/cvx>.
- Gross, D. (2011). Recovering low-rank matrices from few coefficients in any basis. *IEEE Trans. Inf. Theory*, 57(3):1548–1566.
- Gross, D., Liu, Y. K., Flammia, S. T., Becker, S., and Eisert, J. (2010). Quantum state tomography via compressed sensing. *Phys. Rev. Lett.*, 105:150401–150404.
- Gross, D. and Nesme, V. (2010). Note on sampling without replacing from a finite collection of matrices. *arXiv preprint, arXiv:1001.2738*.
- Herman, M. and Strohmer, T. (2009). High-resolution radar via compressed sensing. *IEEE Trans. Signal Process.*, 57:2275–2284.
- Holschneider, C. H. and Berek, J. S. (2000). Ovarian cancer: epidemiology, biology, and prognostic factors. In *Seminars in surgical oncology*, volume 19, pages 3–10. Wiley Online Library.
- Johnstone, I. M. (2001). On the distribution of the largest eigenvalue in principal component analysis. *Ann. Statist.*, 29:295–327.
- Kalbfleisch, J. D. and Prentice, R. L. (2011). *The statistical analysis of failure time data*, volume 360. John Wiley & Sons.
- Keshavan, R. H., Montanari, A., and Oh, S. (2010). Matrix completion from noisy entries. *J. Mach. Learn. Res.*, 11(1):2057–2078.
- Kim, H., Golub, G. H., and Park, H. (2005). Missing value estimation for dna microarray gene expression data: local least squares imputation. *Bioinformatics*, 21(2):187–198.
- Koltchinskii, V. (2011). Von neumann entropy penalization and low-rank matrix estimation. *Ann. Statist.*, 39(6):2936–2973.
- Koltchinskii, V., Lounici, K., and Tsybakov, A. B. (2011). Nuclear-norm penalization and optimal rates for noisy low-rank matrix completion. *Ann. Statist.*, 39:2302–2329.

- Koren, Y., Bell, R., and Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer*, 42:30–37.
- Laurent, B. and Massart, P. (2000). Adaptive estimation of a quadratic functional by model selection. *Ann. Statist.*, 28:1302–1338.
- Li, Y. and Abecasis, G. R. (2006). Mach 1.0: rapid haplotype reconstruction and missing genotype inference. *Am. J. Hum. Genet. S.*, 79(3):2290.
- Lustig, M., Donoho, D. L., Santos, J. M., and Pauly, J. M. (2008). Compressed sensing MRI. *IEEE Signal Process. Mag.*, 27:72–82.
- Mazumder, R., Hastie, T., and Tibshirani, R. (2010). Spectral regularization algorithms for learning large incomplete matrices. *J. Mach. Learn. Res.*, 11:2287–2322.
- Mo, Q. and Li, S. (2011). New bounds on the restricted isometry constant δ_{2k} . *Appl. Comput. Harmon. Anal.*, 31:460–468.
- Mohan, K. and Fazel, M. (2010). New restricted isometry results for noisy low-rank recovery. *Proc. Intl. Sympo. Information Theory (ISIT)*, pages 1573 – 1577.
- Nadler, B. (2010). Nonparametric detection of signals by information theoretic criteria: performance analysis and an improved estimator. *IEEE Trans. Signal Process.*, 58:2746–2756.
- Negahban, S. and Wainwright, M. J. (2011). Estimation of (near) low-rank matrices with noise and high-dimensional scaling. *Ann. Statist.*, 39:1069–1097.
- Oymak, S. and Hassibi, B. (2010). New null space results and recovery thresholds for matrix rank minimization. *arXiv preprint*, arXiv: 1011.6326.
- Oymak, S., Mohan, K., Fazel, M., and Hassibi, B. (2011). A simplified approach to recovery conditions for low-rank matrices. *Proc. Intl. Sympo. Information Theory (ISIT)*, pages 2318–2322.
- Patterson, N., P. A. L. and Reich, D. (2006). Population structure and eigenanalysis. *PLoS genetics*, 2:e190.
- Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics*, 38:904–909.
- Recht, B. (2011). A simpler approach to matrix completion. *J. Mach, Learn. Res.*, 12:3413–3430.
- Recht, B., Fazel, M., and Parrilo, P. (2010). Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.*, 52:471–501.
- Rohde, A. and Tsybakov, A. B. (2011). Estimation of high-dimensional low-rank matrices. *Ann. Statist.*, 39:887–930.
- Salakhutdinov, R. and Srebro, N. (2010). Collaborative filtering in a non-uniform world: Learning with the weighted trace norm. *arXiv preprint arXiv:1002.2780*.
- Scheet, P. and Stephens, M. (2006). A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am. J. Hum. Genet.*, 78(4):629–644.

- Siegel, R., Naishadham, D., and Jemal, A. (2013). Cancer statistics, 2013. *CA-Cancer J. Clin.*, 63(1):11–30.
- Singer, A. and Cucuringu, M. (2010). Uniqueness of low-rank matrix completion by rigidity theory. *SIAM J. Matrix Anal. Appl.*, 31(4):1621–1641.
- Stojnic, M., Xu, W., and Hassibi, B. (2008). Compressed sensing - probabilistic analysis of a null-space characterization. *Acou., Speech and Signal Process, IEEE Internat. Conf. on (ICASSP)*.
- Toh, K.-C. and Yun, S. (2010). An accelerated proximal gradient algorithm for nuclear norm regularized least squares problems. *Pac. J. Optim.*, 6:615–640.
- Tomasi, C. and Kanade, T. (1992). *Shape and motion from image streams: a factorization method parts 2, 8, 10 full report on the orthographic case*. Carnegie Mellon University.
- Tothill, R. W., Tinker, A. V., George, J., Brown, R., Fox, S. B., Lade, S., Johnson, D. S., Trivett, M. K., Etemadmoghadam, D., Locandro, B., et al. (2008). Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome. *Clin. Cancer Res.*, 14(16):5198–5208.
- Tropp, J., Laska, J., Duarte, M., Romberg, J., and Baraniuk, R. (2010). Beyond nyquist: Efficient sampling of sparse, bandlimited signals. *IEEE Trans. Inf. Theory*, 56:520–544.
- Trosset, M. W. (2000). Distance matrix completion by numerical optimization. *Comput. Optim. Appl.*, 17:11–22.
- Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., and Altman, R. B. (2001). Missing value estimation methods for dna microarrays. *Bioinformatics*, 17(6):520–525.
- Vershynin, R. (2011). Spectral norm of products of random and deterministic matrices. *Probab. Theory Relat. Fields*, 150:471–509.
- Wakin, M., Laska, J., Duarte, M., Baron, D., Sarvotham, S., Takhar, D., Kelly, K., and Baraniuk, R. (2006). An architecture for compressive imaging. *Proceedings of the International Conference on Image Processing (ICIP 2006)*, pages 1273–1276.
- Wang, H. and Li, S. (2013). The bounds of restricted isometry constants for low rank matrices recovery. *Sci. China, Ser. A*, 56:1117–1127.
- Wang, X., Li, A., Jiang, Z., and Feng, H. (2006). Missing value estimation for DNA microarray gene expression data by support vector regression imputation and orthogonal coding scheme. *BMC Bioinformatics*, 7(1):32.
- Wang, Y. (2013). Asymptotic equivalence of quantum state tomography and trace regression. *Ann. Statist.*, 41:2462–2504.
- Wax, M. and Kailath, T. (1985). Detection of signals by information theoretic criteria. *IEEE Trans. Acoust., Speech, Signal Processing*, 33:387–392.
- Yu, Z. and Schaid, D. J. (2007). Methods to impute missing genotypes for population data. *Human Genetics*, 122(5):495–504.
- Zhou, S., Kong, L., and Xiu, N. (2013). New bounds for ric in compressed sensing. *J. Oper. Res. Soc. China*, 1:227–237.