

Foundations and Trends® in Machine Learning

Spectral Methods for Data Science: A Statistical Perspective

Suggested Citation: Yuxin Chen, Yuejie Chi, Jianqing Fan and Cong Ma (2021), "Spectral Methods for Data Science: A Statistical Perspective", Foundations and Trends® in Machine Learning: Vol. 14, No. 5, pp 566–806. DOI: 10.1561/2200000079.

Yuxin Chen
Princeton University
yuxin.chen@princeton.edu

Yuejie Chi
Carnegie Mellon University
yuejiechi@cmu.edu

Jianqing Fan
Princeton University
jqfan@princeton.edu

Cong Ma
University of Chicago
congm@uchicago.edu

This article may be used only for the purpose of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval.

now
the essence of knowledge
Boston — Delft

Contents

1	Introduction	568
1.1	Motivating applications	569
1.2	A modern statistical perspective	576
1.3	Organization	577
1.4	What is not here and complementary readings	578
1.5	Notation	579
2	Classical spectral analysis: ℓ_2 perturbation theory	581
2.1	Preliminaries: Basics of matrix analysis	582
2.2	Preliminaries: Distance and angles between subspaces	583
2.3	Perturbation theory for eigenspaces	587
2.4	Perturbation theory for singular subspaces	592
2.5	Eigenvector perturbation for probability transition matrices	595
2.6	Appendix: Proofs of auxiliary lemmas in Section 2.2	599
2.7	Notes	604
3	Applications of ℓ_2 perturbation theory to data science	606
3.1	Preliminaries: Matrix tail bounds	606
3.2	Low-rank matrix denoising	610
3.3	Principal component analysis and factor models	613
3.4	Graph clustering and community recovery	620
3.5	Clustering in Gaussian mixture models	626

3.6	Ranking from pairwise comparisons	642
3.7	Phase retrieval and solving quadratic systems of equations	651
3.8	Matrix completion	663
3.9	Tensor completion	671
3.10	Notes	683
4	Fine-grained analysis: ℓ_∞ and $\ell_{2,\infty}$ perturbation theory	693
4.1	Leave-one-out analysis: An illustrative example	694
4.2	$\ell_{2,\infty}$ eigenspace perturbation under independent noise	701
4.3	$\ell_{2,\infty}$ singular subspace perturbation under independent noise	706
4.4	Application: Entrywise guarantees for matrix completion	708
4.5	Application: Exact community recovery	709
4.6	Distributional theory and uncertainty quantification	718
4.7	Application: Confidence intervals for matrix completion	723
4.8	Appendix A: Proof of Theorem 4.2	729
4.9	Appendix B: Proof of Corollary 4.3	743
4.10	Appendix C: Proof of Theorem 4.4	747
4.11	Appendix D: Proof of Theorem 4.10	750
4.12	Appendix E: Proof of Theorem 4.11	757
4.13	Notes	763
5	Concluding remarks and open problems	766
Acknowledgements		769
References		771

Spectral Methods for Data Science: A Statistical Perspective

Yuxin Chen¹, Yuejie Chi², Jianqing Fan³ and Cong Ma⁴

¹*Princeton University; yuxin.chen@princeton.edu*

²*Carnegie Mellon University; yuejiechi@cmu.edu*

³*Princeton University; jqfan@princeton.edu*

⁴*University of Chicago; congm@uchicago.edu*

ABSTRACT

Spectral methods have emerged as a simple yet surprisingly effective approach for extracting information from massive, noisy and incomplete data. In a nutshell, spectral methods refer to a collection of algorithms built upon the eigenvalues (resp. singular values) and eigenvectors (resp. singular vectors) of some properly designed matrices constructed from data. A diverse array of applications have been found in machine learning, imaging science, financial and econometric modeling, and signal processing, including recommendation systems, community detection, ranking, structured matrix recovery, tensor data estimation, joint shape matching, blind deconvolution, financial investments, risk managements, treatment evaluations, causal inference, amongst others. Due to their simplicity and effectiveness, spectral methods are not only used as a stand-alone estimator, but also frequently employed to facilitate other more sophisticated algorithms to enhance performance.

While the studies of spectral methods can be traced back to classical matrix perturbation theory and the method of

moments, the past decade has witnessed tremendous theoretical advances in demystifying their efficacy through the lens of statistical modeling, with the aid of concentration inequalities and non-asymptotic random matrix theory. This monograph aims to present a systematic, comprehensive, yet accessible introduction to spectral methods from a modern statistical perspective, highlighting their algorithmic implications in diverse large-scale applications. In particular, our exposition gravitates around several central questions that span various applications: how to characterize the sample efficiency of spectral methods in reaching a target level of statistical accuracy, and how to assess their stability in the face of random noise, missing data, and adversarial corruptions? In addition to conventional ℓ_2 perturbation analysis, we present a systematic ℓ_∞ and $\ell_{2,\infty}$ perturbation theory for eigenspace and singular subspaces, which has only recently become available owing to a powerful “leave-one-out” analysis framework.

1

Introduction

In contemporary science and engineering applications, the volume of available data is growing at an enormous rate. The emergence of this trend is due to recent technological advances that have enabled the collection, transmission, storage and processing of data from every corner of our life, in the forms of images, videos, network traffic, email logs, electronic health records, genomic and genetic measurements, high-frequency financial trades, grocery transactions, online exchanges, and so on. In the meantime, modern applications often require reasonings about an unprecedented scale of features or parameters of interest. This gives rise to the pressing demand of developing *low-complexity* algorithms that can effectively distill actionable insights from large-scale and high-dimensional data. In addition to the curse of dimensionality, the challenge is further compounded when the data in hand are noisy, messy, and contain missing features.

Towards addressing the above challenges, *spectral methods* have emerged as a simple yet surprisingly effective approach to information extraction from massive and noisy data. In a nutshell, spectral methods refer to a collection of algorithms built upon the eigenvectors (resp. singular vectors) and eigenvalues (resp. singular values) of some

properly designed matrices generated from data. Remarkably, spectral methods lend themselves to a diverse array of applications in practice, including community detection in networks (Newman, 2006; Abbe, 2017; Rohe *et al.*, 2011; McSherry, 2001), angular synchronization in cryo-EM (Singer and Shkolnisky, 2011; Singer, 2011), joint image alignment (Chen and Candès, 2018), clustering (Von Luxburg, 2007; Ng *et al.*, 2002), ranking (Negahban *et al.*, 2016; Chen and Suh, 2015; Chen *et al.*, 2019b), dimensionality reduction (Belkin and Niyogi, 2003), low-rank matrix estimation (Achlioptas and McSherry, 2007; Keshavan *et al.*, 2010), tensor estimation (Montanari and Sun, 2018; Cai *et al.*, 2019a), covariance and precision matrix estimation (Fan *et al.*, 2013; Fan *et al.*, 2021b), shape reconstruction (Li and Hero, 2004), econometric and financial modeling (Fan *et al.*, 2021a), among others. Motivated by their applicability to numerous real-world problems, this monograph seeks to offer a unified and comprehensive treatment towards establishing the theoretical underpinnings for spectral methods, particularly through a statistical lens.

1.1 Motivating applications

At the heart of spectral methods is the idea that the eigenvectors or singular vectors of certain data matrices reveal crucial information pertaining to the targets of interest. We single out a few examples that epitomize this idea.

Clustering. Clustering corresponds to the grouping of individuals based on their mutual similarities, which constitutes a fundamental task in unsupervised learning and spans numerous applications such as image segmentation (e.g., grouping pixels based on the objects they represent in an image) (Browet *et al.*, 2011) and community detection (e.g., grouping users on the basis of their social circles) (Fortunato and Hric, 2016). For concreteness, let us take a look at a simple scenario with n individuals such that: (1) there exists a latent partitioning that divides all individuals into two groups, with the first $n/2$ individuals belonging to the first group and the rest belonging to the second group (without loss of generality); and (2) we observe pairwise similarity

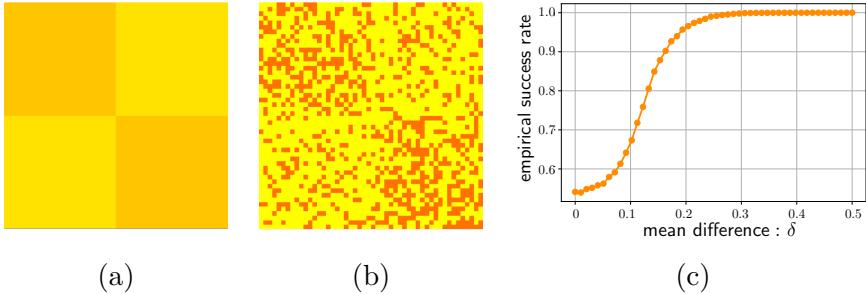


Figure 1.1: Spectral methods for clustering. We plot in (a) an ideal structure of the adjacency matrix \mathbf{A} in (1.1), and in (b) a noisy version which is a realization from the stochastic block model, where $A_{i,j}$ is an independent Bernoulli variable with mean $\frac{1+\delta}{2}$ (resp. $\frac{1-\delta}{2}$) if i and j belong to the same group (resp. different groups). We report in (c) the empirical success rate of the spectral method over 200 Monte Carlo trials in correctly clustering $n = 100$ individuals as the mean difference δ varies.

measurements generated based on their group memberships. Ideally, if we know whether any two individuals belong to the same group or not, then we can form an adjacency matrix $\mathbf{A} = [A_{i,j}]_{1 \leq i,j \leq n}$ such that

$$A_{i,j} = \begin{cases} 1, & \text{if } (i,j) \text{ belongs to the same group,} \\ 0, & \text{else.} \end{cases} \quad (1.1)$$

As a key observation, this matrix \mathbf{A} , as illustrated in Figure 1.1(a), turns out to be a rank-2 matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{1}_{n/2} \mathbf{1}_{n/2}^\top & \\ & \mathbf{1}_{n/2} \mathbf{1}_{n/2}^\top \end{bmatrix} = \frac{1}{2} \mathbf{1}_n \mathbf{1}_n^\top + \frac{1}{2} \begin{bmatrix} \mathbf{1}_{n/2} \\ -\mathbf{1}_{n/2} \end{bmatrix} \begin{bmatrix} \mathbf{1}_{n/2}^\top & -\mathbf{1}_{n/2}^\top \end{bmatrix},$$

where $\mathbf{1}_n$ represents an n -dimensional all-one vector. After subtracting $\frac{1}{2} \mathbf{1}_n \mathbf{1}_n^\top$ from \mathbf{A} , the eigenvector $\mathbf{u}_2 := [\mathbf{1}_{n/2}^\top \ -\mathbf{1}_{n/2}^\top]$ of the remaining component uncovers the underlying group structure; namely, all positive entries of \mathbf{u}_2 represent one group, with all negative entries of \mathbf{u}_2 reflecting another group. In reality, however, we typically only get to collect imprecise information about whether two individuals belong to the same group, thus resulting in a corrupted version of \mathbf{A} (see Figure 1.1(b)). Fortunately, the eigenvector (the one corresponding to \mathbf{u}_2 above) of the observed data matrix (with proper arrangement) might

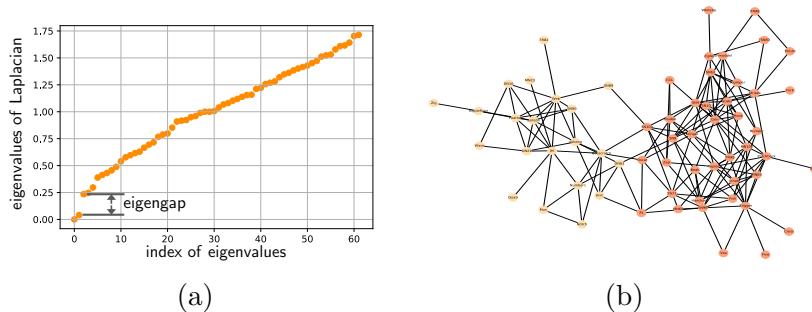


Figure 1.2: Illustration of spectral clustering for 62 dolphins residing in Doubtful Sound, New Zealand. (a) plots the spectrum of the Laplacian matrix of an undirected social network of frequent associations, and (b) illustrates the two communities recovered using the penultimate eigenvector of the Laplacian matrix. Data source: Lusseau *et al.* (2003).

continue to be informative, as long as the noise level is not overly high. To illustrate the practical applicability, we plot in Figure 1.1(c) the numerical performance of this approach, which allows for perfect clustering of all individuals for a wide range of noisy scenarios. Similar ideas continue to fare well on the clustering of real data, where we illustrate in Figure 1.2 that the penultimate eigenvector of a Laplacian matrix (also known as the Fiedler vector) of an undirected social network reveals two communities of 62 dolphins residing in Doubtful Sound, New Zealand.

Principal component analysis (PCA). PCA is arguably one of the most commonly employed tools for data exploration and visualization. Given a collection of data samples $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^p$, PCA seeks to identify a rank- r subspace that explains most of the variability of the data. This is particularly well-grounded when, say, the sample vectors $\{\mathbf{x}_i\}_{1 \leq i \leq n}$ reside primarily within a common rank- r subspace—denoted by \mathbf{U}^* . To extract out this principal subspace, it is instrumental to examine the following sample covariance matrix

$$\mathbf{M} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^\top.$$

If all sample vectors approximately lie within \mathbf{U}^* , then one might be able to infer \mathbf{U}^* by inspecting the rank- r leading eigenspace of \mathbf{M} (or its variants), provided that the signal-to-noise ratio exceeds some reasonable level. This reflects the role of spectral methods in enabling meaningful dimensionality reduction and factor analysis.

In practice, a key benefit of PCA is its ability to remove nuance factors in, and extract out salient features from, each data point. As an illustration, the first four images of Figure 1.3 are representative ones sampled from a face dataset (Georghiades *et al.*, 2001), which correspond to faces of the same person under different illumination and occlusion conditions. In contrast, the “eigenface” (Turk and Pentland, 1991) depicted in the last image of Figure 1.3 corresponds to the first principal component (i.e., $r = 1$), which effectively removes the nuance factors and highlights the feature of the face.



Figure 1.3: Illustration of the eigenface using the Cropped YaleB dataset (Georghiades *et al.*, 2001). The first four images are sampled from this dataset, representing typical images taken under different illumination conditions with various occlusions. The last one represents the eigenface (i.e., the first principal component) of this dataset.

Matrix recovery in the face of missing data. A proliferation of big-data applications has to deal with matrix estimation in the presence of missing data, either due to the infeasibility to acquire complete observations of a massive data matrix (Davenport and Romberg, 2016) such as the Netflix problem in recommender systems (as users only watch and rate a small fraction of movies), or because of the incentive to accelerate computation by means of sub-sampling (Mahoney, 2016). Imagine that we are asked to estimate a large matrix $\mathbf{M}^* = [M_{i,j}^*]_{1 \leq i,j \leq n}$, even though a dominant fraction of its entries are unseen. While in general we cannot predict anything about the missing entries, reliable

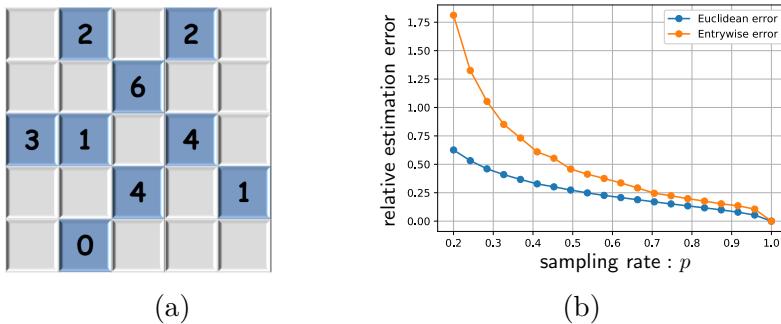


Figure 1.4: Spectral methods for matrix recovery with missing data, where (a) is an illustration of missing data and (b) reports the empirical estimation errors of spectral methods as the sampling rate p varies. Both the relative Euclidean error $\frac{\|\widehat{\mathbf{M}} - \mathbf{M}^*\|_F}{\|\mathbf{M}^*\|_F}$ and the relative entrywise error $\frac{\|\widehat{\mathbf{M}} - \mathbf{M}^*\|_\infty}{\|\mathbf{M}^*\|_\infty}$ are plotted (with $\widehat{\mathbf{M}}$ denoting the matrix estimate and $\|\cdot\|_\infty$ the entrywise ℓ_∞ norm).

estimation might become possible if \mathbf{M}^* is known *a priori* to enjoy a low-rank structure, as is the case in many applications like structure from motion (Tomasi and Kanade, 1992) and sensor network localization (Javanmard and Montanari, 2013). This low-rank assumption motivates the use of spectral methods. More specifically, suppose the entries of \mathbf{M}^* are randomly sampled such that each entry is observed independently with probability $p \in (0, 1]$. An unbiased estimate $\mathbf{M} = [M_{i,j}]_{1 \leq i,j \leq n}$ of \mathbf{M}^* can be readily obtained via rescaling and zero filling (also called the inverse probability weighting method):

$$M_{i,j} = \begin{cases} \frac{1}{p} M_{i,j}^*, & \text{if the } (i, j)\text{-th entry is observed,} \\ 0, & \text{else.} \end{cases}$$

To capture the assumed low-rank structure of \mathbf{M}^* , it is natural to resort to the best rank- r approximation of \mathbf{M} (with r the true rank of \mathbf{M}^*), computable through the rank- r singular value decomposition of \mathbf{M} . Given its (trivial) success when $p = 1$, we expect the algorithm to perform well when p is close to 1. The key question, however, is where the algorithm stands if the vast majority of the entries is missing. While we shall illuminate this in Chapters 3 and 4, Figure 1.4 provides some immediate numerical assessment, which demonstrates the appealing

performance of spectral methods—in terms of both Euclidean and entrywise estimation errors—even when the missing rate is quite high.

Ranking from pairwise comparisons. Another important application of spectral methods arises from the context of ranking, a task of central importance in, say, web search and recommendation systems. In a variety of scenarios, humans find it difficult to simultaneously rank many items, but relatively easier to express pairwise preferences. This gives rise to the problem of ranking based on pairwise comparisons. More specifically, imagine we are given a collection of n items, and wish to identify top-ranked items based on pairwise preferences (with uncertainties in comparison outcomes) between observed pairs of items. A classical statistical model proposed by Bradley and Terry (1952) and Luce (2012) postulates the existence of a set of latent positive scores $\{w_i^*\}_{1 \leq i \leq n}$ —each associated with an item—that determines the ranks of these items. The outcome of the comparison between items i and j is generated in a way that

$$\mathbb{P}(i \text{ beats } j) = \frac{w_i^*}{w_i^* + w_j^*}, \quad 1 \leq i, j \leq n.$$

As it turns out, the preference scores are closely related to the stationary distribution of a Markov chain associated with the above probability kernel, thus forming the basis of spectral ranking algorithms. To elucidate it in a little more detail, let us construct a probability transition matrix $\mathbf{P}^* = [P_{i,j}^*]_{1 \leq i,j \leq n}$ with

$$P_{i,j}^* = \begin{cases} \frac{1}{n} \cdot \frac{w_j^*}{w_i^* + w_j^*}, & \text{if } i \neq j, \\ 1 - \sum_{l:l \neq i} P_{i,l}^*, & \text{if } i = j. \end{cases}$$

Clearly, it forms a probability transition matrix since each element is non-negative and the entries in each row add up to one. It is straightforward to verify that the score vector $\mathbf{w}^* := [w_i^*]_{1 \leq i \leq n}$ satisfies $\mathbf{w}^{*\top} = \mathbf{w}^{*\top} \mathbf{P}^*$, namely \mathbf{w}^* is a left eigenvector of \mathbf{P}^* associated with eigenvalue one. A candidate method then consists of (i) forming an unbiased estimate of \mathbf{P}^* (which can be easily obtained using pairwise comparison outcomes), (ii) computing its left eigenvector (in fact, the leading left eigenvector),

and (iii) reporting the ranking result in accordance with the order of the elements in this eigenvector. This spectral ranking scheme, which shares similar spirit with the celebrated *PageRank* algorithm (Page *et al.*, 1999), exhibits intriguing performance when identifying the top-ranked items, as showcased in the numerical experiments in Figure 1.5(b).

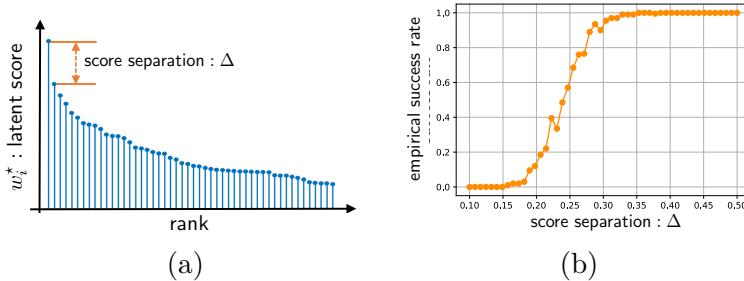


Figure 1.5: Spectral methods for ranking from pairwise comparisons. (a) illustrates the latent preference scores $\{w_i^*\}$ that govern the ranking of items. The empirical success rates in correctly identifying the top-ranked item are plotted in (b) as Δ varies, where Δ represents the separation between the score of the top item and that of the second-ranked item.

A unified theme. In all preceding applications, the core ideas underlying the development of spectral methods can be described in a unified fashion:

1. Identify a key matrix \mathbf{M}^* —which is typically unobserved—whose eigenvectors or singular vectors disclose the information being sought after;
2. Construct a surrogate matrix \mathbf{M} of \mathbf{M}^* using the data samples in hand, and compute the corresponding eigenvectors or singular vectors of this surrogate matrix.

Viewed in this light, this monograph aims to identify key factors—e.g., certain spectral structure of \mathbf{M}^* as well as the size of the approximation error $\mathbf{M} - \mathbf{M}^*$ —that exert main influences on the efficacy of the resultant spectral methods.

1.2 A modern statistical perspective

The idea of spectral methods can be traced back to early statistical literature on methods of moments (e.g., Pearson (1894) and Hansen (1982)), where one seeks to extract key parameters of the probability distributions of interest by examining the empirical moments of data. While classical matrix perturbation theory lays a sensible foundations for the analysis of spectral methods (Stewart and Sun, 1990), the theoretical understanding can be considerably enhanced through the lens of statistical modeling—a way of thinking that has flourished in the past decade. To the best of our knowledge, however, a systematic and comprehensive introduction to the modern statistical foundation of spectral methods, as well as an overview of recent advances, is previously unavailable.

The current monograph aims to fill this gap by developing a coherent and accessible treatment of spectral methods from a modern statistical perspective. Highlighting algorithmic implications that inform practice, our exposition gravitates around the following central questions: how to characterize the sample efficiency of spectral methods in reaching a prescribed accuracy level, and how to assess the stability of spectral methods in the face of random noise, missing data, and adversarial corruptions? We underscore several distinguishing features of our treatment compared to prior studies:

- In comparison to the worst-case performance guarantees derived solely based on classical matrix perturbation theory, our statistical treatment emphasizes the benefit of harnessing the “typical” behavior of data models, which offers key insights into how to harvest performance gains by leveraging intrinsic properties of data generating mechanisms.
- In contrast to classical asymptotic theory (Van der Vaart, 2000), we adopt a non-asymptotic (or finite-sample) analysis framework that draws on tools from recent developments of concentration inequalities (Tropp, 2015) and high-dimensional statistics (Wainwright, 2019). This framework accommodates the scenario where both the sample size and the number of features are enormous, and

unveils a clearer and more complete picture about the interplay and trade-off between salient model parameters.

Another unique feature of this monograph is a principled introduction of *fine-grained entrywise analysis* (e.g., a theory studying ℓ_∞ eigenvector perturbation), which reflects cutting-edge research activities in this area. This is particularly important when, for example, demonstrating the feasibility of exact clustering or perfect ranking in the aforementioned applications. In truth, an effective entrywise analysis framework cannot be readily obtained from classical matrix analysis alone, and has only recently become available owing to the emergence of modern statistical toolboxes. In particular, we shall present a powerful framework, called *leave-one-out analysis*, that proves effective and versatile for delivering fine-grained performance guarantees for spectral methods in a variety of problems.

1.3 Organization

We now present a high-level overview of the structure of this monograph.

- Chapter 2 reviews the fundamentals of classical matrix perturbation theory for spectral analysis, focusing on ℓ_2 -type distances measured by the spectral norm and the Frobenius norm. This chapter covers the celebrated Davis-Kahan $\sin \Theta$ theorem for eigenspace perturbation, the Wedin theorem for singular subspace perturbation, and an extension to probability transition matrices, laying the algebraic foundations for the remaining chapters.
- Chapter 3 explores the utility of ℓ_2 matrix perturbation theory when paired with statistical tools, presenting a unified recipe for statistical analysis empowered by non-asymptotic matrix tail bounds. We develop spectral methods for a variety of statistical data science applications, and derive nearly tight theoretical guarantees (up to logarithmic factors) based on this unified recipe.
- Chapter 4 develops fine-grained perturbation theory for spectral analysis in terms of ℓ_∞ and $\ell_{2,\infty}$ metrics, based on a leave-one-out analysis framework rooted in probability theory. Its effectiveness

is demonstrated through concrete applications including community recovery and matrix completion. This analysis framework also enables a non-asymptotic distributional theory for spectral methods, which paves the way for uncertainty quantification in applications like noisy matrix completion.

- Chapter 5 concludes this monograph by identifying a few directions that are worthy of future investigation.

While this monograph pursues a coherent and accessible treatment that might appeal to a broad audience, it does not necessarily deliver the sharpest possible results for the applications discussed herein in terms of the logarithmic terms and/or pre-constants. The bibliographic notes at the end of each chapter contain information about the state-of-the-art theory for each application as a pointer to further readings.

1.4 What is not here and complementary readings

The topics presented in this monograph do not cover the tensor decomposition methods studied in another recent strand of work (Anandkumar *et al.*, 2014). While such tensor-based methods are also sometimes referred to as spectral methods, their primary focus is to invoke tensor decomposition to learn latent variables, based on higher-order moments estimated from data samples. We elect not to discuss this class of methods but instead refer the interested reader to the recently published monograph by Janzamin *et al.* (2019). Another monograph by Kannan and Vempala (2009) provides an in-depth computational and algorithmic treatment of spectral methods from the perspective of theoretical computer science. The applications and results covered therein (e.g., fast matrix multiplication) complement the ones presented in the current monograph. In addition, spectral methods have been frequently employed to initialize nonconvex optimization algorithms. We will not elaborate on the nonconvex optimization aspect here but instead recommend the reader to the recent overview article by Chi *et al.* (2019). Finally, spectral methods are widely adopted to estimate high-dimensional covariance and precision matrices, and extract latent factors for econometric and

statistical modeling. This topic alone has a huge literature, and we refer the interested reader to Fan *et al.* (2020b) for in-depth discussions.

1.5 Notation

Before moving forward, let us introduce some notation that will be used throughout this monograph.

First of all, we reserve boldfaced symbols for vectors, matrices and tensors. For any matrix \mathbf{A} , let $\sigma_j(\mathbf{A})$ (resp. $\lambda_j(\mathbf{A})$) represent its j -th largest singular value (resp. eigenvalue). In particular, $\sigma_{\max}(\mathbf{A})$ (resp. $\lambda_{\max}(\mathbf{A})$) stands for the largest singular value (resp. eigenvalue) of \mathbf{A} , while $\sigma_{\min}(\mathbf{A})$ (resp. $\lambda_{\min}(\mathbf{A})$) indicates the smallest singular value (resp. eigenvalue) of \mathbf{A} . We use \mathbf{A}^\top to denote the transpose of \mathbf{A} , and let $\mathbf{A}_{i,\cdot}$ and $\mathbf{A}_{\cdot,i}$ indicate the i -th row and the i -th column of \mathbf{A} , respectively. We follow standard conventions by letting \mathbf{I}_n be the $n \times n$ identity matrix, $\mathbf{1}_n$ the n -dimensional all-one vector, and $\mathbf{0}_n$ the n -dimensional all-zero vector; we shall often suppress the subscript as long as it is clear from the context. The i -th standard basis vector is denoted by \mathbf{e}_i throughout. The notation $\mathcal{O}^{n \times r}$ ($r \leq n$) represents the set of all $n \times r$ orthonormal matrices (whose columns are orthonormal). Moreover, we refer to $[n]$ as the set $\{1, \dots, n\}$.

Next, we turn to vector and matrix norms. For any vector \mathbf{v} , we denote by $\|\mathbf{v}\|_2$, $\|\mathbf{v}\|_1$ and $\|\mathbf{v}\|_\infty$ its ℓ_2 norm, ℓ_1 norm and ℓ_∞ norm, respectively. For any matrix $\mathbf{A} = [A_{i,j}]_{1 \leq i \leq m, 1 \leq j \leq n}$, we let $\|\mathbf{A}\|$, $\|\mathbf{A}\|_*$, $\|\mathbf{A}\|_F$ and $\|\mathbf{A}\|_\infty$ represent respectively its spectral norm (i.e., the largest singular value of \mathbf{A}), its nuclear norm (i.e., the sum of singular values of \mathbf{A}), its Frobenius norm (i.e., $\|\mathbf{A}\|_F := \sqrt{\sum_{i,j} A_{i,j}^2}$), and its entrywise ℓ_∞ norm (i.e., $\|\mathbf{A}\|_\infty := \max_{i,j} |A_{i,j}|$). We also refer to $\|\mathbf{A}\|_{2,\infty}$ as the $\ell_{2,\infty}$ norm of \mathbf{A} , defined as $\|\mathbf{A}\|_{2,\infty} := \max_i \|\mathbf{A}_{i,\cdot}\|_2$. Similarly, we define the $\ell_{\infty,2}$ norm of \mathbf{A} as $\|\mathbf{A}\|_{\infty,2} := \|\mathbf{A}^\top\|_{2,\infty}$. In addition, for any matrices $\mathbf{A} = [A_{i,j}]_{1 \leq i \leq m, 1 \leq j \leq n}$ and $\mathbf{B} = [B_{i,j}]_{1 \leq i \leq m, 1 \leq j \leq n}$, the inner product of \mathbf{A} and \mathbf{B} is defined as and denoted by $\langle \mathbf{A}, \mathbf{B} \rangle = \sum_{1 \leq i \leq m, 1 \leq j \leq n} A_{i,j} B_{i,j} = \text{Tr}(\mathbf{A}^\top \mathbf{B})$.

When it comes to diagonal matrices, we employ $\text{diag}([\theta_1, \theta_2, \dots, \theta_r])$ to abbreviate the diagonal matrix with diagonal elements $\theta_1, \dots, \theta_r$. For any diagonal matrix $\Theta = \text{diag}([\theta_1, \theta_2, \dots, \theta_r])$, we adopt the shorthand

notation $\sin \Theta := \text{diag}([\sin \theta_1, \sin \theta_2, \dots, \sin \theta_r])$; the notation $\sin^2 \Theta$, $\cos \Theta$, and $\cos^2 \Theta$ is defined analogously.

Finally, this monograph makes heavy use of the following standard notation: (1) $f(n) = O(g(n))$ or $f(n) \lesssim g(n)$ means that there exists a universal constant $c > 0$ such that $|f(n)| \leq c|g(n)|$ holds for all sufficiently large n ; (2) $f(n) \gtrsim g(n)$ means that there exists a universal constant $c > 0$ such that $|f(n)| \geq c|g(n)|$ holds for all sufficiently large n ; (3) $f(n) \asymp g(n)$ means that there exist universal constants $c_1, c_2 > 0$ such that $c_1|g(n)| \leq |f(n)| \leq c_2|g(n)|$ holds for all sufficiently large n ; and (4) $f(n) = o(g(n))$ indicates that $f(n)/g(n) \rightarrow 0$ as $n \rightarrow \infty$. Additionally, we sometimes use $f(n) \gg g(n)$ (resp. $f(n) \ll g(n)$) to indicate that there exists some sufficiently large (resp. small) universal constant $c > 0$ such that $|f(n)| \geq c|g(n)|$ (resp. $|f(n)| \leq c|g(n)|$).

2

Classical spectral analysis: ℓ_2 perturbation theory

Characterizing the performance of spectral methods requires understanding the perturbation of eigenspaces and/or that of singular subspaces. Classical matrix perturbation theory (e.g., Stewart and Sun (1990)) offers elementary toolkits that prove effective for this purpose, which we review in this chapter.

Setting the stage, consider a real-valued matrix \mathbf{M}^* and its perturbed version as follows

$$\mathbf{M} = \mathbf{M}^* + \mathbf{E}, \tag{2.1}$$

where $\mathbf{E} = \mathbf{M} - \mathbf{M}^*$ denotes a real-valued perturbation or error matrix. In statistical applications, \mathbf{M} can be an observed or estimated data matrix such as the sample covariance matrix, and \mathbf{M}^* is the target matrix such as the population covariance matrix. This chapter primarily aims to address the following questions by means of elementary linear algebra:

1. For a symmetric matrix \mathbf{M}^* , how does the eigenspace change in response to a symmetric perturbation matrix \mathbf{E} ?
2. For a general matrix \mathbf{M}^* , how is the singular subspace affected as a result of the perturbation matrix \mathbf{E} ?

We shall also explore eigenvector perturbation for a special class of asymmetric matrices: probability transition matrices.

2.1 Preliminaries: Basics of matrix analysis

We begin this chapter by gathering a few elementary materials in matrix analysis that prove useful for our theoretical development. The readers familiar with matrix analysis can proceed directly to Section 2.2.

Unitarily invariant norms. Among all matrix norms, the family of unitarily invariant norms defined below is of central interest, which subsumes as special cases the spectral norm $\|\cdot\|$ and the Frobenius norm $\|\cdot\|_F$.

Definition 2.1. A matrix norm $\|\cdot\|$ on $\mathbb{R}^{m \times n}$ is said to be unitarily invariant if

$$\|A\| = \|U^\top A V\|$$

holds for any matrix $A \in \mathbb{R}^{m \times n}$ and any two square orthonormal matrices $U \in \mathcal{O}^{m \times m}$ and $V \in \mathcal{O}^{n \times n}$.

This class of matrix norms enjoys several useful properties, as summarized in the following lemma. The proof can be found in Stewart and Sun (1990, Theorem 3.9).

Lemma 2.1. For any unitarily invariant norm $\|\cdot\|$, one has

$$\begin{aligned} \|AB\| &\leq \|A\| \cdot \|B\|, & \|AB\| &\leq \|B\| \cdot \|A\|, \\ \|AB\| &\geq \|A\| \sigma_{\min}(B), & \|AB\| &\geq \|B\| \sigma_{\min}(A). \end{aligned}$$

Perturbation bounds for eigenvalues and singular values. Next, we review classical perturbation bounds for eigenvalues of symmetric matrices and for singular values of general matrices.

Lemma 2.2 (Weyl's inequality for eigenvalues). Let $A, E \in \mathbb{R}^{n \times n}$ be two real symmetric matrices. For every $1 \leq i \leq n$, the i -th largest eigenvalues of A and $A + E$ obey

$$|\lambda_i(A) - \lambda_i(A + E)| \leq \|E\|. \quad (2.2)$$

Proof. See Equation (1.63) in Tao (2012). \square

Lemma 2.3 (Weyl's inequality for singular values). Let $\mathbf{A}, \mathbf{E} \in \mathbb{R}^{m \times n}$ be two general matrices. Then for every $1 \leq i \leq \min\{m, n\}$, the i -th largest singular values of \mathbf{A} and $\mathbf{A} + \mathbf{E}$ obey

$$|\sigma_i(\mathbf{A} + \mathbf{E}) - \sigma_i(\mathbf{A})| \leq \|\mathbf{E}\|.$$

Proof. See Exercise 1.3.22 in Tao (2012). \square

An immediate implication of Lemma 2.2 (resp. Lemma 2.3) is that the eigenvalues of a real symmetric matrix (resp. the singular values of a general matrix) are stable vis-à-vis small perturbations.

2.2 Preliminaries: Distance and angles between subspaces

In order to develop perturbation theory for eigenspaces and singular subspaces, we first need to delineate a metric that quantifies the proximity of two subspaces in a meaningful way.

2.2.1 Setup and notation

Consider two r -dimensional subspaces \mathcal{U}^* and \mathcal{U} in \mathbb{R}^n , where $1 \leq r \leq n$. One can represent these two subspaces by two matrices $\mathbf{U}^* \in \mathbb{R}^{n \times r}$ and $\mathbf{U} \in \mathbb{R}^{n \times r}$, whose columns form an orthonormal basis of \mathcal{U}^* and \mathcal{U} , respectively. Here and throughout, we shall use \mathcal{U} and its matrix representation \mathbf{U} interchangeably whenever it is clear from the context.

For the sake of convenience, we further introduce two $n \times (n - r)$ matrices \mathbf{U}_\perp^* and \mathbf{U}_\perp , such that $[\mathbf{U}^*, \mathbf{U}_\perp^*]$ and $[\mathbf{U}, \mathbf{U}_\perp]$ are both $n \times n$ orthonormal matrices. In other words, \mathbf{U}_\perp^* and \mathbf{U}_\perp represent the orthogonal complement of \mathbf{U}^* and \mathbf{U} , respectively.

2.2.2 Distance metrics and principal angles

Global rotational ambiguity. To measure the distance between the two subspaces \mathcal{U} and \mathcal{U}^* , a naive idea is to employ the “metric” $\|\mathbf{U} - \mathbf{U}^*\|$, where $\|\cdot\|$ is a certain norm of interest (e.g., the spectral norm or the Frobenius norm). An immediate drawback arises, however, since this

“metric” does not take into account the global rotational ambiguity—namely, for any rotation matrix $\mathbf{R} \in \mathcal{O}^{r \times r}$, the columns of the matrix $\mathbf{U}\mathbf{R}$ also form a valid orthonormal basis of \mathcal{U} . This means that even when the two subspaces \mathcal{U} and \mathcal{U}^* coincide, one might still have $\|\mathbf{U} - \mathbf{U}^*\| \neq 0$, depending on how we rotate these matrices.

Valid choices of distance and angles. The takeaway of the above discussion is that any meaningful metric employed to measure the proximity of two subspaces should account for the rotational ambiguity properly. In what follows, we single out a few widely used metrics that meet such a requirement.

1. *Distance with optimal rotation.* Given the global rotational ambiguity, it is natural to first adjust the rotation matrix suitably before computing the distance. One choice is to measure the distance upon optimal rotation, namely,

$$\text{dist}_{\|\cdot\|}(\mathbf{U}, \mathbf{U}^*) := \min_{\mathbf{R} \in \mathcal{O}^{r \times r}} \|\mathbf{U}\mathbf{R} - \mathbf{U}^*\|, \quad (2.3)$$

where $\|\cdot\|$ is a certain norm to be chosen (e.g., the spectral norm or the Frobenius norm).

2. *Distance between projection matrices.* As an established fact, the projection matrix onto a subspace \mathcal{U} —given by $\mathbf{U}\mathbf{U}^\top$ —is unique and unaffected by how \mathbf{U} is rotated (since $\mathbf{U}\mathbf{U}^\top = \mathbf{U}\mathbf{R}\mathbf{R}^\top\mathbf{U}^\top$ for any rotation matrix $\mathbf{R} \in \mathcal{O}^{r \times r}$). The rotational invariance of the projection matrix motivates us to define the distance between \mathcal{U} and \mathcal{U}^* as follows

$$\text{dist}_{\mathbf{p}, \|\cdot\|}(\mathbf{U}, \mathbf{U}^*) := \|\mathbf{U}\mathbf{U}^\top - \mathbf{U}^*\mathbf{U}^{*\top}\|, \quad (2.4)$$

where, as usual, $\|\cdot\|$ is a certain matrix norm of interest, and the subscript \mathbf{p} stands for projection.

3. *Geometric construction via principal angles.* Let $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq 0$ be the singular values of $\mathbf{U}^\top\mathbf{U}^*$, arranged in descending order. Given that $\|\mathbf{U}^\top\mathbf{U}^*\| \leq \|\mathbf{U}\| \|\mathbf{U}^*\| = 1$, all the singular values $\{\sigma_i\}_{i=1}^r$ fall within the interval $[0, 1]$. Therefore, one can

define the principal angles (or canonical angles) between the two subspaces of interest as

$$\theta_i := \arccos(\sigma_i) \quad \text{for all } 1 \leq i \leq r, \quad (2.5)$$

which clearly satisfy

$$0 \leq \theta_1 \leq \dots \leq \theta_r \leq \pi/2. \quad (2.6)$$

To see why this definition makes sense, consider the simplest example where $r = 1$. In this case, the principal angle θ_1 coincides with the conventionally defined angle between two unit vectors \mathbf{U} and \mathbf{U}^* . Armed with these angles, one might measure the distance between the subspaces \mathcal{U} and \mathcal{U}^* through the following metric

$$\text{dist}_{\sin, \|\cdot\|}(\mathbf{U}, \mathbf{U}^*) := \|\sin \Theta\|, \quad (2.7)$$

where $\|\cdot\|$ is again some matrix norm to be selected, and

$$\Theta := \begin{bmatrix} \theta_1 & & \\ & \ddots & \\ & & \theta_r \end{bmatrix}, \quad \sin \Theta := \begin{bmatrix} \sin \theta_1 & & \\ & \ddots & \\ & & \sin \theta_r \end{bmatrix}. \quad (2.8)$$

With slight abuse of notation, we can define other diagonal matrices such as $\cos \Theta$ analogously, where $\cos(\cdot)$ is applied in an entrywise manner to the diagonal elements of Θ . Such matrices will be useful for future discussions.

2.2.3 Intimate connections between the distance metrics

It turns out that the metrics (2.3), (2.4) and (2.7) introduced above are tightly related, as we shall explain in this subsection. The proofs of all the results in this subsection are deferred to Section 2.6.

To begin with, we take a look at the relation between $\text{dist}_{\mathbf{p}, \|\cdot\|}(\cdot, \cdot)$ and $\text{dist}_{\sin, \|\cdot\|}(\cdot, \cdot)$, which is perhaps best illuminated by the following lemma.

Lemma 2.4. Consider the settings of Section 2.2.1. If $2r \leq n$, then the singular values of $\mathbf{U}\mathbf{U}^\top - \mathbf{U}^*\mathbf{U}^{*\top}$ (including zeros) are given by

$$\underbrace{\sin \theta_r, \sin \theta_r, \sin \theta_{r-1}, \sin \theta_{r-1}, \dots, \sin \theta_1, \sin \theta_1}_{2r}, \underbrace{0, 0, \dots, 0}_{n-2r}.$$

In a nutshell, Lemma 2.4 establishes an explicit link between (a) the difference of the projection matrices and (b) the principal angles between the two subspaces of interest. This lemma and its analysis unveil the following crucial equivalence relation under two of our favorite norms—the spectral norm and the Frobenius norm; in light of this, we might refer to these metrics as the $\sin \Theta$ distances from time to time.

Lemma 2.5. Consider the settings of Section 2.2.1, and recall the definition of $\sin \Theta$ in (2.8). For any $1 \leq r \leq n$, one has

$$\|UU^\top - U^*U^{*\top}\| = \|\sin \Theta\| = \|U_\perp^\top U^*\| = \|U^\top U_\perp^*\|; \quad (2.9a)$$

$$\frac{1}{\sqrt{2}}\|UU^\top - U^*U^{*\top}\|_F = \|\sin \Theta\|_F = \|U_\perp^\top U^*\|_F = \|U^\top U_\perp^*\|_F. \quad (2.9b)$$

Next, we move on to demonstrate the (near) equivalence of $\text{dist}_{\|\cdot\|}(\cdot, \cdot)$ and $\text{dist}_{p, \|\cdot\|}(\cdot, \cdot)$ under the above-mentioned two norms.

Lemma 2.6. Under the settings of Section 2.2.1, for any $1 \leq r \leq n$, one has¹

$$\begin{aligned} \|UU^\top - U^*U^{*\top}\| &\leq \min_{R \in \mathcal{O}^{r \times r}} \|UR - U^*\| \leq \sqrt{2}\|UU^\top - U^*U^{*\top}\|; \\ \frac{1}{\sqrt{2}}\|UU^\top - U^*U^{*\top}\|_F &\leq \min_{R \in \mathcal{O}^{r \times r}} \|UR - U^*\|_F \leq \|UU^\top - U^*U^{*\top}\|_F. \end{aligned}$$

In words, $\text{dist}_{\|\cdot\|}(\cdot, \cdot)$ and $\text{dist}_{p, \|\cdot\|}(\cdot, \cdot)$ are equivalent up to a factor of $\sqrt{2}$, when $\|\cdot\|$ is the spectral norm or the Frobenius norm.

2.2.4 The distance metrics of choice in this monograph

In conclusion, the following metrics, which are seemingly distinct at first glance, are (nearly) equivalent in measuring the distance between two subspaces U and U^* :

- 1) $\|UU^\top - U^*U^{*\top}\|$
- 2) $\|\sin \Theta\|$
- 3) $\|U_\perp^\top U^*\| = \|U^\top U_\perp^*\|$
- 4) $\min_{R \in \mathcal{O}^{r \times r}} \|UR - U^*\|$

¹It is straightforward to verify that the upper bounds on both $\min_{R \in \mathcal{O}^{r \times r}} \|UR - U^*\|$ and $\min_{R \in \mathcal{O}^{r \times r}} \|UR - U^*\|_F$ are attainable when $U = [1, 0]^\top$ and $U^* = [0, 1]^\top$.

when $\|\cdot\|$ represents either the spectral norm or the Frobenius norm. Viewed in this light, we shall mainly concentrate on the following metrics throughout the rest of this monograph:

$$\text{dist}(\mathbf{U}, \mathbf{U}^*) := \min_{\mathbf{R} \in \mathcal{O}^{r \times r}} \|\mathbf{U}\mathbf{R} - \mathbf{U}^*\|; \quad (2.10a)$$

$$\text{dist}_F(\mathbf{U}, \mathbf{U}^*) := \min_{\mathbf{R} \in \mathcal{O}^{r \times r}} \|\mathbf{U}\mathbf{R} - \mathbf{U}^*\|_F. \quad (2.10b)$$

2.3 Perturbation theory for eigenspaces

Armed with the above metrics for subspace distances, we are in a position to identify key factors that affect the perturbation of eigenvectors and eigenspaces.

2.3.1 Setup and notation

Let \mathbf{M}^* and $\mathbf{M} = \mathbf{M}^* + \mathbf{E}$ be two $n \times n$ real symmetric matrices. We express the eigendecomposition of \mathbf{M}^* and \mathbf{M} as follows

$$\mathbf{M}^* = \sum_{i=1}^n \lambda_i^* \mathbf{u}_i^* \mathbf{u}_i^{*\top} = \begin{bmatrix} \mathbf{U}^* & \mathbf{U}_\perp^* \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_\perp^* \end{bmatrix} \begin{bmatrix} \mathbf{U}^{*\top} \\ \mathbf{U}_\perp^{*\top} \end{bmatrix}; \quad (2.11)$$

$$\mathbf{M} = \sum_{i=1}^n \lambda_i \mathbf{u}_i \mathbf{u}_i^\top = \begin{bmatrix} \mathbf{U} & \mathbf{U}_\perp \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_\perp \end{bmatrix} \begin{bmatrix} \mathbf{U}^\top \\ \mathbf{U}_\perp^\top \end{bmatrix}. \quad (2.12)$$

Here, $\{\lambda_i\}$ (resp. $\{\lambda_i^*\}$) denote the eigenvalues of \mathbf{M} (resp. \mathbf{M}^*), and \mathbf{u}_i (resp. \mathbf{u}_i^*) stands for the eigenvector associated with the eigenvalue λ_i (resp. λ_i^*). Additionally, we take

$$\mathbf{U} := [\mathbf{u}_1, \dots, \mathbf{u}_r] \in \mathbb{R}^{n \times r}, \quad \mathbf{U}_\perp := [\mathbf{u}_{r+1}, \dots, \mathbf{u}_n] \in \mathbb{R}^{n \times (n-r)},$$

$$\mathbf{\Lambda} := \text{diag}([\lambda_1, \dots, \lambda_r]), \quad \mathbf{\Lambda}_\perp := \text{diag}([\lambda_{r+1}, \dots, \lambda_n]).$$

The matrices \mathbf{U}^* , \mathbf{U}_\perp^* , $\mathbf{\Lambda}^*$, and $\mathbf{\Lambda}_\perp^*$ are defined analogously.

2.3.2 A warm-up example

In general, the eigenvector/eigenspace of a real symmetric matrix might change drastically even upon a small perturbation. To understand this,

consider the following toy example borrowed from Hsu (2016):

$$\mathbf{M}^* = \begin{bmatrix} 1 + \epsilon & 0 \\ 0 & 1 - \epsilon \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} -\epsilon & \epsilon \\ \epsilon & \epsilon \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} 1 & \epsilon \\ \epsilon & 1 \end{bmatrix},$$

where $0 < \epsilon < 1$ can be arbitrarily small. It is straightforward to check that the leading eigenvectors of \mathbf{M}^* and \mathbf{M} are given respectively by

$$\mathbf{u}_1^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Consequently, we have

$$\|\mathbf{u}_1 \mathbf{u}_1^\top - \mathbf{u}_1^* \mathbf{u}_1^{*\top}\| = \frac{1}{\sqrt{2}}, \quad \text{and} \quad \|\mathbf{u}_1 \mathbf{u}_1^\top - \mathbf{u}_1^* \mathbf{u}_1^{*\top}\|_{\text{F}} = 1, \quad (2.13)$$

which are both quite large regardless of the size of ϵ or the size of the perturbation $\|\mathbf{E}\|$.

On closer inspection, this “pathological” behavior comes up due to the fact that perturbation size ϵ is comparable to the eigengap of \mathbf{M}^* (namely, $\lambda_1(\mathbf{M}^*) - \lambda_2(\mathbf{M}^*) = 2\epsilon$). This hints at the important role played by the eigengap in influencing eigenspace perturbation.

2.3.3 The Davis-Kahan $\sin\Theta$ theorem

At the core of classical eigenspace perturbation theory lies the landmark result of Davis and Kahan (1970), which delivers powerful eigenspace perturbation bounds in terms of the size of the perturbation matrix as well as the associated eigengap. Here and throughout, for any symmetric matrix \mathbf{A} , we denote by $\text{eigenvalues}(\mathbf{A})$ the set of eigenvalues of \mathbf{A} .

Theorem 2.7 (Davis-Kahan’s $\sin\Theta$ theorem). Consider the settings in Section 2.3.1. Assume that

$$\text{eigenvalues}(\mathbf{A}^*) \subseteq [\alpha, \beta], \quad (2.14a)$$

$$\text{eigenvalues}(\mathbf{A}_\perp) \subseteq (-\infty, \alpha - \Delta] \cup [\beta + \Delta, \infty) \quad (2.14b)$$

for some quantities $\alpha, \beta \in \mathbb{R}$ and eigengap $\Delta > 0$. Then one has

$$\text{dist}(\mathbf{U}, \mathbf{U}^*) \leq \sqrt{2} \|\sin \Theta\| \leq \frac{\sqrt{2} \|\mathbf{E} \mathbf{U}^*\|}{\Delta} \leq \frac{\sqrt{2} \|\mathbf{E}\|}{\Delta}; \quad (2.15a)$$

$$\text{dist}_F(\mathbf{U}, \mathbf{U}^*) \leq \sqrt{2} \|\sin \Theta\|_F \leq \frac{\sqrt{2} \|\mathbf{E} \mathbf{U}^*\|_F}{\Delta} \leq \frac{\sqrt{2r} \|\mathbf{E}\|}{\Delta}. \quad (2.15b)$$

This conclusion remains valid if Assumption (2.14) is replaced by

$$\text{eigenvalues}(\Lambda^*) \subseteq (-\infty, \alpha - \Delta] \cup [\beta + \Delta, \infty); \quad (2.16a)$$

$$\text{eigenvalues}(\Lambda_{\perp}) \subseteq [\alpha, \beta]. \quad (2.16b)$$

Remark 2.1. In fact, Theorem 2.7 can be generalized to accommodate any unitarily invariant norm $\|\cdot\|$, in the sense that

$$\|\sin \Theta\| \leq \frac{\|\mathbf{E} \mathbf{U}^*\|}{\Delta}. \quad (2.17)$$

The proof of Theorem 2.7 and Remark 2.1 is quite elementary and can be found in Section 2.3.4.

Remark 2.2. As we shall demonstrate in Chapter 4, the above bounds that involve $\|\mathbf{E} \mathbf{U}^*\|$ and $\|\mathbf{E} \mathbf{U}^*\|_F$ are particularly useful when \mathbf{E} exhibits special structure (e.g., row sparsity or column sparsity).

Theorem 2.7 is commonly referred to as the Davis-Kahan $\sin \Theta$ theorem, given that it concerns the $\sin \Theta$ distance between subspaces. Both bounds scale linearly with the perturbation size, and are inversely proportional to the eigengap Δ . Informally, if we view $\|\mathbf{E}\|$ as the noise size and interpret the eigengap as the “signal strength” (which dictates how easy it is to distinguish the r eigenvalues of interest from the remaining spectrum), then Theorem 2.7 asserts that the eigenspace perturbation degrades gracefully as the signal-to-noise-ratio decreases.

The careful reader might notice that Theorem 2.7 stays silent on the allowable size $\|\mathbf{E}\|$ of the perturbation. Note, however, that a restriction on $\|\mathbf{E}\|$ is somewhat hidden in Assumptions (2.14) and (2.16). When the eigenvalues in Λ^* (resp. Λ) and Λ_{\perp}^* (resp. Λ_{\perp}) are suitably ordered, it is oftentimes more convenient to work with the following corollary, which makes apparent the constraint on the size $\|\mathbf{E}\|$ with regard to the eigengap of \mathbf{M}^* .

Corollary 2.8. Consider the settings in Section 2.3.1. Suppose that $|\lambda_1^*| \geq |\lambda_2^*| \geq \cdots \geq |\lambda_r^*| > |\lambda_{r+1}^*| \geq \cdots \geq |\lambda_n^*|$ and $|\lambda_1| \geq |\lambda_2| \geq$

$\dots \geq |\lambda_n|$ (i.e., the eigenvalues are sorted by their magnitudes). If $\|\mathbf{E}\| < (1 - 1/\sqrt{2})(|\lambda_r^*| - |\lambda_{r+1}^*|)$, then

$$\text{dist}(\mathbf{U}, \mathbf{U}^*) \leq \sqrt{2} \|\sin \Theta\| \leq \frac{2\|\mathbf{E}\|}{|\lambda_r^*| - |\lambda_{r+1}^*|} \leq \frac{2\|\mathbf{E}\|}{|\lambda_r^*| - |\lambda_{r+1}^*|}; \quad (2.18a)$$

$$\text{dist}_F(\mathbf{U}, \mathbf{U}^*) \leq \sqrt{2} \|\sin \Theta\|_F \leq \frac{2\|\mathbf{E}\|_F}{|\lambda_r^*| - |\lambda_{r+1}^*|} \leq \frac{2\sqrt{r}\|\mathbf{E}\|}{|\lambda_r^*| - |\lambda_{r+1}^*|}. \quad (2.18b)$$

The proof of Corollary 2.8 is also given in Section 2.3.4.

2.3.4 Proof of the Davis-Kahan $\sin\Theta$ theorem

Proof of Theorem 2.7. The proof proceeds by controlling the distance metric $\|\mathbf{U}_\perp^\top \mathbf{U}^*\|$, where $\|\cdot\|$ denotes a unitarily invariant norm.

We start by proving the theorem under Assumption (2.14), and claim that it suffices to consider the case where

$$\alpha = -\beta \leq 0. \quad (2.19)$$

In fact, if this condition is violated, then one can employ a “centering” trick by enforcing global offset to \mathbf{M}^* and \mathbf{M} as follows

$$\mathbf{M}_c^* = \mathbf{M}^* - \frac{\alpha + \beta}{2} \mathbf{I}_n, \quad \text{and} \quad \mathbf{M}_c = \mathbf{M} - \frac{\alpha + \beta}{2} \mathbf{I}_n.$$

It is straightforwardly seen that (a) \mathbf{M}_c^* (resp. \mathbf{M}_c) and \mathbf{M}^* (resp. \mathbf{M}) share the same eigenvectors; (b) the eigenvalues of \mathbf{M}_c^* (resp. \mathbf{M}_c) associated with \mathbf{U}^* (resp. \mathbf{U}_\perp) reside within $[-\gamma, \gamma]$ (resp. $(-\infty, -\gamma - \Delta] \cup [\gamma + \Delta, \infty)$), where $\gamma = \frac{\beta - \alpha}{2} \geq 0$. Consequently, this reduces to a scenario that resembles (2.19). In addition, we isolate two immediate consequences of Assumptions (2.14) and (2.19) that prove useful:

$$\|\mathbf{\Lambda}^*\| \leq \beta, \quad \text{and} \quad \sigma_{\min}(\mathbf{\Lambda}_\perp) \geq \beta + \Delta, \quad (2.20)$$

where we recall that $\sigma_{\min}(\mathbf{\Lambda}_\perp)$ is the minimal singular value of $\mathbf{\Lambda}_\perp$.

Armed with the above spectral conditions, we are prepared to study $\mathbf{U}_\perp^\top \mathbf{U}^*$. This is controlled through the following identity (obtained by the definition of eigenvectors):

$$\mathbf{U}_\perp^\top (\mathbf{M} - \mathbf{M}^*) \mathbf{U}^* = \mathbf{\Lambda}_\perp \mathbf{U}_\perp^\top \mathbf{U}^* - \mathbf{U}_\perp^\top \mathbf{U}^* \mathbf{\Lambda}^*, \quad (2.21)$$

Let $\mathbf{R} := (\mathbf{M} - \mathbf{M}^*) \mathbf{U}^* = \mathbf{E} \mathbf{U}^*$. The triangle inequality then tells us that

$$\begin{aligned} \|\mathbf{U}_\perp^\top \mathbf{R}\| &\geq \|\Lambda_\perp \mathbf{U}_\perp^\top \mathbf{U}^*\| - \|\mathbf{U}_\perp^\top \mathbf{U}^* \Lambda^*\| \\ &\geq \sigma_{\min}(\Lambda_\perp) \|\mathbf{U}_\perp^\top \mathbf{U}^*\| - \|\Lambda^*\| \cdot \|\mathbf{U}_\perp^\top \mathbf{U}^*\| \\ &\geq (\beta + \Delta - \beta) \|\mathbf{U}_\perp^\top \mathbf{U}^*\| = \Delta \|\mathbf{U}_\perp^\top \mathbf{U}^*\|. \end{aligned} \quad (2.22)$$

Here, the middle line follows from Lemma 2.1 in Section 2.1, whereas the last inequality arises from the properties (2.20). As a consequence,

$$\|\mathbf{U}_\perp^\top \mathbf{U}^*\| \leq \frac{\|\mathbf{U}_\perp^\top \mathbf{R}\|}{\Delta} \leq \frac{\|\mathbf{R}\|}{\Delta} = \frac{\|\mathbf{E} \mathbf{U}^*\|}{\Delta},$$

where the second inequality follows again from Lemma 2.1 and $\|\mathbf{U}_\perp\| = 1$. When $\|\cdot\|$ is either the spectral norm or the Frobenius norm, combining the preceding inequality with Lemmas 2.5-2.6 and the facts $\|\mathbf{U}^*\| = 1$ and $\|\mathbf{U}^*\|_F = \sqrt{r}$ immediately establishes the theorem for this case.

Next, we turn to the scenario where Assumption (2.16) is in effect; it can be analyzed in a similar manner and hence we remark only on the difference. Assuming (2.19) holds without loss of generality, we have

$$\|\Lambda_\perp\| \leq \beta, \quad \text{and} \quad \sigma_{\min}(\Lambda^*) \geq \beta + \Delta. \quad (2.23)$$

Applying the triangle inequality to (2.21) in a different way yields

$$\begin{aligned} \|\mathbf{U}_\perp^\top \mathbf{R}\| &\geq \|\mathbf{U}_\perp^\top \mathbf{U}^* \Lambda^*\| - \|\Lambda_\perp \mathbf{U}_\perp^\top \mathbf{U}^*\| \\ &\geq \sigma_{\min}(\Lambda^*) \|\mathbf{U}_\perp^\top \mathbf{U}^*\| - \|\Lambda_\perp\| \cdot \|\mathbf{U}_\perp^\top \mathbf{U}^*\| \\ &\geq (\beta + \Delta - \beta) \|\mathbf{U}_\perp^\top \mathbf{U}^*\| = \Delta \|\mathbf{U}_\perp^\top \mathbf{U}^*\|, \end{aligned}$$

a conclusion that coincides with (2.22). The rest of the proof is the same as the one in the previous case.

Before concluding, we remark that $\|\mathbf{U}_\perp^\top \mathbf{U}^*\| = \|\sin \Theta\|$ holds for any unitarily invariant norm $\|\cdot\|$; see Li (1998, Lemma 2.1). This together with the above analysis leads to Remark 2.1.

Proof of Corollary 2.8. We first examine the spectral ranges of Λ_\perp and Λ^* . Let $\lambda_i(\mathbf{M}^*)$ (resp. $\lambda_i(\mathbf{M})$) be the i -th largest eigenvalue of \mathbf{M}^* (resp. \mathbf{M}), sorted by their values (as opposed to their magnitudes). Then Weyl's inequality (cf. Lemma 2.2 in Section 2.1) asserts that

$$|\lambda_i(\mathbf{M}) - \lambda_i(\mathbf{M}^*)| \leq \|\mathbf{E}\|, \quad 1 \leq i \leq n.$$

Suppose that \mathbf{M}^* has r_1 positive (resp. $r_2 = r - r_1$ negative) eigenvalues whose magnitudes exceed $|\lambda_{r+1}^*|$. Then for any i obeying $1 \leq i \leq r_1$ or $i > n - r_2$, the triangle inequality gives

$$\begin{aligned} |\lambda_i(\mathbf{M})| &\geq |\lambda_i(\mathbf{M}^*)| - \|\mathbf{E}\| > |\lambda_r^*| - (1 - 1/\sqrt{2})(|\lambda_r^*| - |\lambda_{r+1}^*|) \\ &> |\lambda_{r+1}^*| + (1 - 1/\sqrt{2})(|\lambda_r^*| - |\lambda_{r+1}^*|) \geq |\lambda_{r+1}^*| + \|\mathbf{E}\|, \end{aligned}$$

where the last inequality arises from our assumption on $\|\mathbf{E}\|$. On the contrary, if $r_1 < i \leq n - r_2$, then one has

$$|\lambda_i(\mathbf{M})| \leq |\lambda_{r+1}^*| + \|\mathbf{E}\|.$$

As a consequence, there are exactly r (resp. $n - r$) eigenvalues of \mathbf{M} whose magnitudes exceed (resp. lie below) $|\lambda_{r+1}^*| + \|\mathbf{E}\|$.

The above observation together with the ordering $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$ implies

$$\text{eigenvalues}(\mathbf{\Lambda}_\perp) \subseteq [-|\lambda_{r+1}^*| - \|\mathbf{E}\|, |\lambda_{r+1}^*| + \|\mathbf{E}\|].$$

In addition, the assumption that $|\lambda_1^*| \geq |\lambda_2^*| \geq \dots \geq |\lambda_n^*|$ tells us that

$$\text{eigenvalues}(\mathbf{\Lambda}^*) \subseteq (-\infty, -|\lambda_r^*|] \cup [|\lambda_r^*|, \infty).$$

Taking $\beta = -\alpha = |\lambda_{r+1}^*| + \|\mathbf{E}\|$ and $\Delta = |\lambda_r^*| - |\lambda_{r+1}^*| - \|\mathbf{E}\| > (|\lambda_r^*| - |\lambda_{r+1}^*|)/\sqrt{2}$, we can invoke Theorem 2.7 under Assumption (2.16) to establish the advertised results.

2.4 Perturbation theory for singular subspaces

There is no shortage of scenarios where the data matrices under consideration are asymmetric or rectangular. In these cases, one is often asked to study singular value decomposition (SVD) rather than eigen-decomposition. Fortunately, the eigenspace perturbation theory can be naturally extended to accommodate perturbation of singular subspaces.

2.4.1 Setup and notation

Let \mathbf{M}^* and $\mathbf{M} = \mathbf{M}^* + \mathbf{E}$ be two matrices in $\mathbb{R}^{n_1 \times n_2}$ (without loss of generality, we assume $n_1 \leq n_2$), whose SVDs are given respectively by

$$\mathbf{M}^* = \sum_{i=1}^{n_1} \sigma_i^* \mathbf{u}_i^* \mathbf{v}_i^{*\top} = \begin{bmatrix} \mathbf{U}^* & \mathbf{U}_\perp^* \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}^* & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_\perp^* & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}^{*\top} \\ \mathbf{V}_\perp^{*\top} \end{bmatrix}; \quad (2.24)$$

$$\mathbf{M} = \sum_{i=1}^{n_1} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top = \begin{bmatrix} \mathbf{U} & \mathbf{U}_\perp \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_\perp & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}^\top \\ \mathbf{V}_\perp^\top \end{bmatrix}. \quad (2.25)$$

Here, $\sigma_1 \geq \dots \geq \sigma_{n_1}$ (resp. $\sigma_1^* \geq \dots \geq \sigma_{n_1}^*$) stand for the singular values of \mathbf{M} (resp. \mathbf{M}^*) arranged in descending order, \mathbf{u}_i (resp. \mathbf{u}_i^*) denotes the left singular vector associated with the singular value σ_i (resp. σ_i^*), and \mathbf{v}_i (resp. \mathbf{v}_i^*) represents the right singular vector associated with σ_i (resp. σ_i^*). In addition, we denote

$$\begin{aligned} \boldsymbol{\Sigma} &:= \text{diag}([\sigma_1, \dots, \sigma_r]), & \boldsymbol{\Sigma}_\perp &:= \text{diag}([\sigma_{r+1}, \dots, \sigma_{n_1}]), \\ \mathbf{U} &:= [\mathbf{u}_1, \dots, \mathbf{u}_r] \in \mathbb{R}^{n_1 \times r}, & \mathbf{U}_\perp &:= [\mathbf{u}_{r+1}, \dots, \mathbf{u}_{n_1}] \in \mathbb{R}^{n_1 \times (n_1-r)}, \\ \mathbf{V} &:= [\mathbf{v}_1, \dots, \mathbf{v}_r] \in \mathbb{R}^{n_2 \times r}, & \mathbf{V}_\perp &:= [\mathbf{v}_{r+1}, \dots, \mathbf{v}_{n_2}] \in \mathbb{R}^{n_2 \times (n_2-r)}. \end{aligned}$$

The matrices $\boldsymbol{\Sigma}^*, \boldsymbol{\Sigma}_\perp^*, \mathbf{U}^*, \mathbf{U}_\perp^*, \mathbf{V}^*, \mathbf{V}_\perp^*$ are defined analogously.

2.4.2 Wedin's $\sin\Theta$ theorem

Wedin (1972) developed a perturbation bound for singular subspaces that parallels the Davis-Kahan $\sin\Theta$ theorem for eigenspaces. In what follows, we present a version that is convenient for subsequent discussions in this monograph.

Theorem 2.9 (Wedin's $\sin\Theta$ theorem). Consider the settings in Section 2.4.1. If $\|\mathbf{E}\| < \sigma_r^* - \sigma_{r+1}^*$, then one has

$$\begin{aligned} \max \{ \text{dist}(\mathbf{U}, \mathbf{U}^*), \text{dist}(\mathbf{V}, \mathbf{V}^*) \} &\leq \frac{\sqrt{2} \max \{ \|\mathbf{E}^\top \mathbf{U}^*\|, \|\mathbf{E} \mathbf{V}^*\| \}}{\sigma_r^* - \sigma_{r+1}^* - \|\mathbf{E}\|}; \\ \max \{ \text{dist}_F(\mathbf{U}, \mathbf{U}^*), \text{dist}_F(\mathbf{V}, \mathbf{V}^*) \} &\leq \frac{\sqrt{2} \max \{ \|\mathbf{E}^\top \mathbf{U}^*\|_F, \|\mathbf{E} \mathbf{V}^*\|_F \}}{\sigma_r^* - \sigma_{r+1}^* - \|\mathbf{E}\|}. \end{aligned}$$

This theorem simultaneously controls the perturbation of left and right singular subspaces. As a worthy note, both the interaction between \mathbf{E} and \mathbf{U}^* , and that between \mathbf{E} and \mathbf{V}^* , come into play in determining the perturbation bounds. In particular, if $\|\mathbf{E}\| < (1 - 1/\sqrt{2})(\sigma_r^* - \sigma_{r+1}^*)$, then one can apply Lemma 2.1 in Section 2.1 to obtain

$$\max \{ \text{dist}(\mathbf{U}, \mathbf{U}^*), \text{dist}(\mathbf{V}, \mathbf{V}^*) \} \leq \frac{2\|\mathbf{E}\|}{\sigma_r^* - \sigma_{r+1}^*}, \quad (2.26a)$$

$$\max \{ \text{dist}_F(\mathbf{U}, \mathbf{U}^*), \text{dist}_F(\mathbf{V}, \mathbf{V}^*) \} \leq \frac{2\sqrt{r}\|\mathbf{E}\|}{\sigma_r^* - \sigma_{r+1}^*}, \quad (2.26b)$$

akin to the eigenspace perturbation bounds (2.18).

2.4.3 Proof of the Wedin $\sin\Theta$ theorem

We now present a proof of the Wedin theorem. Similar to the proof of the Davis-Kahan theorem, we start by bounding $\|\mathbf{U}_\perp^\top \mathbf{U}^*\|$, where $\|\cdot\|$ stands for any unitarily invariant norm. To this end, it is seen that

$$\begin{aligned} \mathbf{U}_\perp^\top \mathbf{U}^* &= \mathbf{U}_\perp^\top (\mathbf{U}^* \Sigma^* \mathbf{V}^{*\top}) \mathbf{V}^* \Sigma^{*-1} \\ &= \mathbf{U}_\perp^\top (\mathbf{M} - \mathbf{E} - \mathbf{U}_\perp^* \Sigma_\perp^* \mathbf{V}_\perp^{*\top}) \mathbf{V}^* \Sigma^{*-1} \\ &= \mathbf{U}_\perp^\top (\mathbf{U} \Sigma \mathbf{V}^\top + \mathbf{U}_\perp \Sigma_\perp \mathbf{V}_\perp^\top - \mathbf{E} - \mathbf{U}_\perp^* \Sigma_\perp^* \mathbf{V}_\perp^{*\top}) \mathbf{V}^* \Sigma^{*-1} \\ &= \Sigma_\perp \mathbf{V}_\perp^\top \mathbf{V}^* \Sigma^{*-1} - \mathbf{U}_\perp^\top \mathbf{E} \mathbf{V}^* \Sigma^{*-1}. \end{aligned} \quad (2.27)$$

Here, the first identity is valid as long as Σ^* is invertible (which is guaranteed since $\sigma_{\min}(\Sigma^*) = \sigma_r^* > \sigma_{r+1}^* + \|\mathbf{E}\| > 0$ under our assumption), the second line follows from the identities $\mathbf{M} - \mathbf{E} = \mathbf{M}^*$ and $\mathbf{M}^* = \mathbf{U}^* \Sigma^* \mathbf{V}^{*\top} + \mathbf{U}_\perp^* \Sigma_\perp^* \mathbf{V}_\perp^{*\top}$, the third line holds since $\mathbf{M} = \mathbf{U} \Sigma \mathbf{V}^\top + \mathbf{U}_\perp \Sigma_\perp \mathbf{V}_\perp^\top$, whereas the last identity exploits the property

$$\mathbf{U}_\perp^\top \mathbf{U} = \mathbf{0}, \quad \text{and} \quad \mathbf{V}_\perp^{*\top} \mathbf{V}^* = \mathbf{0}.$$

Applying the triangle inequality and Lemma 2.1 in Section 2.1 to the identity (2.27) yields

$$\begin{aligned} \|\mathbf{U}_\perp^\top \mathbf{U}^*\| &\leq \|\Sigma_\perp\| \cdot \|\mathbf{V}_\perp^\top \mathbf{V}^*\| \cdot \|\Sigma^{*-1}\| + \|\mathbf{U}_\perp^\top\| \cdot \|\mathbf{E} \mathbf{V}^*\| \cdot \|\Sigma^{*-1}\| \\ &= \sigma_{r+1} \cdot \|\mathbf{V}_\perp^\top \mathbf{V}^*\| \cdot \frac{1}{\sigma_r^*} + \|\mathbf{E} \mathbf{V}^*\| \cdot \frac{1}{\sigma_r^*} \\ &\leq \frac{\sigma_{r+1}^* + \|\mathbf{E}\|}{\sigma_r^*} \|\mathbf{V}_\perp^\top \mathbf{V}^*\| + \frac{\|\mathbf{E} \mathbf{V}^*\|}{\sigma_r^*}. \end{aligned} \quad (2.28)$$

Here, the second line uses the properties $\|\Sigma^{*-1}\| = 1/\sigma_r^*$ and $\|\Sigma_\perp\| = \sigma_{r+1}$, while the last inequality follows from Weyl's inequality $\sigma_{r+1} \leq \sigma_{r+1}^* + \|\mathbf{E}\|$ (cf. Lemma 2.3 in Section 2.1). Repeating the same argument yields

$$\|\mathbf{V}_\perp^\top \mathbf{V}^*\| \leq \frac{\|\mathbf{E}^\top \mathbf{U}^*\|}{\sigma_r^*} + \frac{\sigma_{r+1}^* + \|\mathbf{E}\|}{\sigma_r^*} \|\mathbf{U}_\perp^\top \mathbf{U}^*\|. \quad (2.29)$$

To finish up, combine the inequalities (2.28) and (2.29) to obtain

$$\begin{aligned} \max \{ \| \mathbf{U}_\perp^\top \mathbf{U}^* \|, \| \mathbf{V}_\perp^\top \mathbf{V}^* \| \} &\leq \frac{\max \{ \| \mathbf{E}^\top \mathbf{U}^* \|, \| \mathbf{E} \mathbf{V}^* \| \}}{\sigma_r^*} \\ &\quad + \frac{\sigma_{r+1}^* + \| \mathbf{E} \|}{\sigma_r^*} \max \{ \| \mathbf{U}_\perp^\top \mathbf{U}^* \|, \| \mathbf{V}_\perp^\top \mathbf{V}^* \| \}. \end{aligned}$$

When $\| \mathbf{E} \| < \sigma_r^* - \sigma_{r+1}^*$, we can rearrange terms to arrive at

$$\max \{ \| \mathbf{U}_\perp^\top \mathbf{U}^* \|, \| \mathbf{V}_\perp^\top \mathbf{V}^* \| \} \leq \frac{\max \{ \| \mathbf{E}^\top \mathbf{U}^* \|, \| \mathbf{E} \mathbf{V}^* \| \}}{\sigma_r^* - \sigma_{r+1}^* - \| \mathbf{E} \|}.$$

The proof is then completed by invoking Lemmas 2.5 and 2.6.

2.5 Eigenvector perturbation for probability transition matrices

Thus far, our eigenvector perturbation analysis has been constrained to the set of *symmetric* matrices. Note, however, that the utility of eigenvectors is by no means confined to symmetric matrices. In fact, eigenvector analysis plays a vital role in studying asymmetric matrices as well, most notably the family of probability transition matrices of Markov chains. This section explores how to extend eigenvector perturbation theory to accommodate an important class of probability transition matrices associated with *reversible* Markov chains.

2.5.1 Background, setup and notation

Before presenting the formulation, we remind the readers that a matrix $\mathbf{P} \in \mathbb{R}^{n \times n}$ is a probability transition matrix if it is composed of non-negative entries with each row summing to 1, which is used to describe the state transition of a Markov chain over a set of n states. Of special interest is the stationary distribution of \mathbf{P} , denoted by a probability vector $\boldsymbol{\pi} = [\pi_i]_{1 \leq i \leq n}$, that satisfies

$$\boldsymbol{\pi} \geq \mathbf{0}, \quad \mathbf{1}^\top \boldsymbol{\pi} = 1, \quad \text{and} \quad \boldsymbol{\pi}^\top \mathbf{P} = \boldsymbol{\pi}^\top. \quad (2.30)$$

In words, the distribution $\boldsymbol{\pi}$ is invariant with respect to \mathbf{P} . Clearly, $\boldsymbol{\pi}$ is the left eigenvector of \mathbf{P} associated with eigenvalue 1, with the corresponding right eigenvector given by $\mathbf{1}$. By the Gershgorin circle

theorem (see, e.g., Olver *et al.* (2006)), the modulus of all eigenvalues must be bounded by the maximum of the row sum, which is 1. Given that 1 is an eigenvalue of \mathbf{P} , the largest modulus of the eigenvalues of \mathbf{P} is precisely 1, and therefore $\boldsymbol{\pi}$ is the *leading* left eigenvector of \mathbf{P} . In addition, a Markov chain is said to be reversible when the following *detailed balance* equations are satisfied:

$$\pi_i P_{i,j} = \pi_j P_{j,i}, \quad \text{for all } 1 \leq i, j \leq n, \quad (2.31)$$

where $\boldsymbol{\pi} = [\pi_i]_{1 \leq i \leq n}$ is the stationary distribution obeying (2.30). It will be seen in the proof of Theorem 2.10 that all eigenvalues of such a matrix \mathbf{P} are real. For readers who wish an introduction to the basics of Markov chains, we recommend the monograph by Brémaud (2013).

In this section, we consider the probability transition matrix $\mathbf{P}^* \in \mathbb{R}^{n \times n}$ of a reversible Markov chain, as well as its perturbed version—also in the form of a probability transition matrix:

$$\mathbf{P} = \mathbf{P}^* + \mathbf{E} \in \mathbb{R}^{n \times n}.$$

The leading left eigenvectors of \mathbf{P}^* and \mathbf{P} —or equivalently, the vectors representing their stationary distributions—are denoted by $\boldsymbol{\pi}^*$ and $\boldsymbol{\pi}$, respectively. Here, we allow \mathbf{E} to be fairly general, meaning that \mathbf{P} does not necessarily represent a reversible Markov chain. The question is: how does the matrix \mathbf{E} affect the perturbation $\boldsymbol{\pi} - \boldsymbol{\pi}^*$ of the leading left eigenvector of interest?

Additionally, we find it helpful to introduce several notation frequently used in the studies of Markov chains. Instead of operating under the usual ℓ_2 norm, the stationary distribution $\boldsymbol{\pi}$ equips us with a new set of norms. Specifically, for a strictly positive probability vector $\boldsymbol{\pi} = [\pi_i]_{1 \leq i \leq n}$, any vector $\mathbf{x} = [x_i]_{1 \leq i \leq n}$ and any matrix \mathbf{A} , it is useful to introduce the vector norm $\|\mathbf{x}\|_{\boldsymbol{\pi}} := \sqrt{\sum_i \pi_i x_i^2}$ and the corresponding matrix norm $\|\mathbf{A}\|_{\boldsymbol{\pi}} := \sup_{\|\mathbf{x}\|_{\boldsymbol{\pi}}=1} \|\mathbf{A}\mathbf{x}\|_{\boldsymbol{\pi}}$.

2.5.2 Perturbation of the leading eigenvector

Now we are ready to present the perturbation bound for the leading left eigenvector of a probability transition matrix, a result originally developed in Chen *et al.* (2019b).

Theorem 2.10. Consider the settings in Section 2.5.1. Suppose that \mathbf{P}^* represents a reversible Markov chain, whose stationary distribution vector $\boldsymbol{\pi}^*$ is strictly positive. Assume that

$$\|\mathbf{E}\|_{\boldsymbol{\pi}^*} < 1 - \max \{ \lambda_2(\mathbf{P}^*), -\lambda_n(\mathbf{P}^*) \}. \quad (2.32)$$

Then one has

$$\|\boldsymbol{\pi} - \boldsymbol{\pi}^*\|_{\boldsymbol{\pi}^*} \leq \frac{\|\boldsymbol{\pi}^{*\top} \mathbf{E}\|_{\boldsymbol{\pi}^*}}{1 - \max \{ \lambda_2(\mathbf{P}^*), -\lambda_n(\mathbf{P}^*) \} - \|\mathbf{E}\|_{\boldsymbol{\pi}^*}}.$$

The similarity between Theorem 2.10 and Corollary 2.8 is noteworthy. Indeed, recalling that the largest eigenvalue of the probability transition matrix \mathbf{P}^* is precisely 1, one might view $1 - \max \{ \lambda_2(\mathbf{P}^*), -\lambda_n(\mathbf{P}^*) \}$ as the gap between the first and the second largest eigenvalues of \mathbf{P}^* (in magnitude), akin to the eigengap $|\lambda_r^*| - |\lambda_{r+1}^*|$ in Corollary 2.8 (with $r = 1$). In words, Theorem 2.10 guarantees that as long as the size of the perturbation matrix \mathbf{E} is not too large, the perturbation of the leading left eigenvector—or equivalently, the perturbation of the stationary distribution of the associated Markov chain—is proportional to the size of the noise when projected onto the direction $\boldsymbol{\pi}^*$, as measured by $\|\boldsymbol{\pi}^{*\top} \mathbf{E}\|_{\boldsymbol{\pi}^*}$. As we shall demonstrate in Section 3.6, this perturbation theory delivers powerful techniques for analyzing the ranking problem described previously in Chapter 1.

Remark 2.3. Sensitivity and perturbation analyses for the steady-state distributions of Markov chains have been studied in the literature; see, e.g., Mitrophanov (2005), Liu (2012), Jiang *et al.* (2017), and Rudolf and Schweizer (2018) and the references therein.

2.5.3 Proof of Theorem 2.10

Since $\boldsymbol{\pi}^*$ and $\boldsymbol{\pi}$ denote respectively the leading left eigenvectors of \mathbf{P}^* and \mathbf{P} , namely,

$$\boldsymbol{\pi}^{*\top} \mathbf{P}^* = \boldsymbol{\pi}^{*\top}, \quad \text{and} \quad \boldsymbol{\pi}^{\top} \mathbf{P} = \boldsymbol{\pi}^{\top},$$

the perturbation $\boldsymbol{\pi} - \boldsymbol{\pi}^*$ admits the following decomposition

$$\boldsymbol{\pi}^{\top} - \boldsymbol{\pi}^{*\top} = \boldsymbol{\pi}^{\top} \mathbf{P} - \boldsymbol{\pi}^{*\top} \mathbf{P}^* = (\boldsymbol{\pi} - \boldsymbol{\pi}^*)^{\top} \mathbf{P} + \boldsymbol{\pi}^{*\top} (\mathbf{P} - \mathbf{P}^*)$$

$$\begin{aligned}
&= (\boldsymbol{\pi} - \boldsymbol{\pi}^*)^\top (\mathbf{P} - \mathbf{P}^*) + (\boldsymbol{\pi} - \boldsymbol{\pi}^*)^\top \mathbf{P}^* + \boldsymbol{\pi}^{*\top} (\mathbf{P} - \mathbf{P}^*) \\
&= (\boldsymbol{\pi} - \boldsymbol{\pi}^*)^\top (\mathbf{P} - \mathbf{P}^*) + (\boldsymbol{\pi} - \boldsymbol{\pi}^*)^\top (\mathbf{P}^* - \mathbf{1}\boldsymbol{\pi}^{*\top}) + \boldsymbol{\pi}^{*\top} (\mathbf{P} - \mathbf{P}^*).
\end{aligned}$$

Here, the last relation hinges upon the fact that $\boldsymbol{\pi}$ and $\boldsymbol{\pi}^*$ are probability vectors, and hence $(\boldsymbol{\pi} - \boldsymbol{\pi}^*)^\top \mathbf{1} = 1 - 1 = 0$. Apply the triangle inequality with respect to the norm $\|\cdot\|_{\boldsymbol{\pi}^*}$ to obtain

$$\begin{aligned}
\|\boldsymbol{\pi} - \boldsymbol{\pi}^*\|_{\boldsymbol{\pi}^*} &\leq \|(\boldsymbol{\pi} - \boldsymbol{\pi}^*)^\top (\mathbf{P} - \mathbf{P}^*)\|_{\boldsymbol{\pi}^*} + \|(\boldsymbol{\pi} - \boldsymbol{\pi}^*)^\top (\mathbf{P}^* - \mathbf{1}\boldsymbol{\pi}^{*\top})\|_{\boldsymbol{\pi}^*} \\
&\quad + \|\boldsymbol{\pi}^{*\top} (\mathbf{P} - \mathbf{P}^*)\|_{\boldsymbol{\pi}^*} \\
&\leq \left(\|\mathbf{P} - \mathbf{P}^*\|_{\boldsymbol{\pi}^*} + \|\mathbf{P}^* - \mathbf{1}\boldsymbol{\pi}^{*\top}\|_{\boldsymbol{\pi}^*} \right) \|\boldsymbol{\pi} - \boldsymbol{\pi}^*\|_{\boldsymbol{\pi}^*} \\
&\quad + \|\boldsymbol{\pi}^{*\top} (\mathbf{P} - \mathbf{P}^*)\|_{\boldsymbol{\pi}^*},
\end{aligned}$$

where the last line relies on the definition of the matrix norm $\|\cdot\|_{\boldsymbol{\pi}^*}$. Rearranging terms, we are left with

$$\|\boldsymbol{\pi} - \boldsymbol{\pi}^*\|_{\boldsymbol{\pi}^*} \leq \frac{\|\boldsymbol{\pi}^{*\top} (\mathbf{P} - \mathbf{P}^*)\|_{\boldsymbol{\pi}^*}}{1 - \|\mathbf{P} - \mathbf{P}^*\|_{\boldsymbol{\pi}^*} - \|\mathbf{P}^* - \mathbf{1}\boldsymbol{\pi}^{*\top}\|_{\boldsymbol{\pi}^*}},$$

with the proviso that $\|\mathbf{P} - \mathbf{P}^*\|_{\boldsymbol{\pi}^*} + \|\mathbf{P}^* - \mathbf{1}\boldsymbol{\pi}^{*\top}\|_{\boldsymbol{\pi}^*} < 1$. The proof would then be completed as long as one could justify that

$$\|\mathbf{P}^* - \mathbf{1}\boldsymbol{\pi}^{*\top}\|_{\boldsymbol{\pi}^*} = \max \{ \lambda_2(\mathbf{P}^*), -\lambda_n(\mathbf{P}^*) \}. \quad (2.33)$$

Proof of the identity (2.33). Let $\boldsymbol{\pi}^* = [\pi_i^*]_{1 \leq i \leq n}$, and define a diagonal matrix $\boldsymbol{\Pi}^* = \text{diag}([\pi_1^*, \dots, \pi_n^*]) \in \mathbb{R}^{n \times n}$. From the definition of the norm $\|\cdot\|_{\boldsymbol{\pi}^*}$ (both the matrix version and the vector version), it is easily seen that for any matrix \mathbf{A} ,

$$\begin{aligned}
\|\mathbf{A}\|_{\boldsymbol{\pi}^*} &= \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_{\boldsymbol{\pi}^*}}{\|\mathbf{x}\|_{\boldsymbol{\pi}^*}} = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|(\boldsymbol{\Pi}^*)^{1/2} \mathbf{A}(\boldsymbol{\Pi}^*)^{-1/2} (\boldsymbol{\Pi}^*)^{1/2} \mathbf{x}\|_2}{\|(\boldsymbol{\Pi}^*)^{1/2} \mathbf{x}\|_2} \\
&= \sup_{\mathbf{v} \neq \mathbf{0}} \frac{\|(\boldsymbol{\Pi}^*)^{1/2} \mathbf{A}(\boldsymbol{\Pi}^*)^{-1/2} \mathbf{v}\|_2}{\|\mathbf{v}\|_2} = \|(\boldsymbol{\Pi}^*)^{1/2} \mathbf{A}(\boldsymbol{\Pi}^*)^{-1/2}\|, \quad (2.34)
\end{aligned}$$

with $\|\cdot\|$ the usual spectral norm, where the last line replaces $(\boldsymbol{\Pi}^*)^{1/2} \mathbf{x}$ with \mathbf{v} . Consequently, we obtain

$$\|\mathbf{P}^* - \mathbf{1}\boldsymbol{\pi}^{*\top}\|_{\boldsymbol{\pi}^*} = \|(\boldsymbol{\Pi}^*)^{1/2} (\mathbf{P}^* - \mathbf{1}\boldsymbol{\pi}^{*\top}) (\boldsymbol{\Pi}^*)^{-1/2}\|$$

$$= \|S^* - \pi_{1/2}^*(\pi_{1/2}^*)^\top\|,$$

where we define $S^* := (\Pi^*)^{1/2} P^* (\Pi^*)^{-1/2}$ and $\pi_{1/2}^* := [\sqrt{\pi_i^*}]_{1 \leq i \leq n}$. Several basic properties regarding S^* are in order; see Brémaud (2013, Chapter 6.2).

- (a) Since P^* represents a reversible Markov chain with stationary distribution π^* , the matrix S^* is symmetric, whose eigenvalues are real-valued. This can be verified by the detailed balance equations (2.31).
- (b) Given that S^* is obtained via a similarity transformation of P^* , we see that S^* and P^* share the same set of eigenvalues. This can easily be verified from the definition of eigenvectors:

$$S^* \xi = \lambda \xi \iff P^* (\Pi^*)^{-1/2} \xi = \lambda (\Pi^*)^{-1/2} \xi.$$

- (c) In particular, $\lambda_1(S^*) = \lambda_1(P^*) = 1$, and $\pi_{1/2}^*$ is precisely the eigenvector of S^* associated with $\lambda_1(S^*) = 1$. Thus, from the eigendecomposition of the symmetric matrix S^* , it is easy to see that the eigenvalues of $S^* - \pi_{1/2}^*(\pi_{1/2}^*)^\top$ are $0, \lambda_2(S^*), \dots, \lambda_n(S^*)$.

Taking the preceding facts collectively, we reach

$$\begin{aligned} \|S^* - \pi_{1/2}^*(\pi_{1/2}^*)^\top\| &\stackrel{(i)}{=} \max \{|\lambda_2(S^*)|, |\lambda_n(S^*)|\} \\ &= \max \{\lambda_2(S^*), -\lambda_n(S^*)\} \stackrel{(ii)}{=} \max \{\lambda_2(P^*), -\lambda_n(P^*)\}. \end{aligned}$$

Here, (i) relies on Property (c), while (ii) follows from Property (b). This concludes the proof. \square

2.6 Appendix: Proofs of auxiliary lemmas in Section 2.2

2.6.1 Proof of Lemma 2.4

Given that singular values are unitarily invariant, it suffices to look at the singular values of the following matrix

$$\begin{bmatrix} \mathbf{U}^\top \\ \mathbf{U}_\perp^\top \end{bmatrix} (\mathbf{U} \mathbf{U}^\top - \mathbf{U}^* \mathbf{U}^{*\top}) [\mathbf{U}_\perp^*, \mathbf{U}^*] = \begin{bmatrix} \mathbf{U}^\top \mathbf{U}_\perp^* & \mathbf{0} \\ \mathbf{0} & -\mathbf{U}_\perp^\top \mathbf{U}^* \end{bmatrix}. \quad (2.35)$$

Consequently, the singular values of $\mathbf{U}\mathbf{U}^\top - \mathbf{U}^*\mathbf{U}^{*\top}$ are composed of those of $\mathbf{U}^\top\mathbf{U}_\perp^*$ and those of $\mathbf{U}_\perp^*\mathbf{U}^*$ combined. It then boils down to characterizing the spectrum of $\mathbf{U}^\top\mathbf{U}_\perp^*$ and $\mathbf{U}_\perp^*\mathbf{U}^*$.

To pin down the singular values of $\mathbf{U}^\top\mathbf{U}_\perp^*$, we first turn attention to the eigenvalues of $\mathbf{U}^\top\mathbf{U}_\perp^*\mathbf{U}_\perp^*\top\mathbf{U}$. Assuming that the SVD of $\mathbf{U}^\top\mathbf{U}^*$ is given by $\mathbf{X}\mathbf{\Sigma}\mathbf{Y}^\top$ (where \mathbf{X} and \mathbf{Y} are $r \times r$ orthonormal matrices, and $\mathbf{\Sigma}$ is diagonal), we can derive

$$\begin{aligned}\mathbf{U}^\top\mathbf{U}_\perp^*\mathbf{U}_\perp^*\top\mathbf{U} &= \mathbf{U}^\top(\mathbf{I}_n - \mathbf{U}^*\mathbf{U}^{*\top})\mathbf{U} = \mathbf{U}^\top\mathbf{U} - \mathbf{U}^\top\mathbf{U}^*\mathbf{U}^{*\top}\mathbf{U} \\ &= \mathbf{I}_r - \mathbf{X}\mathbf{\Sigma}^2\mathbf{X}^\top = \mathbf{X}(\mathbf{I}_r - \cos^2\Theta)\mathbf{X}^\top \\ &= \mathbf{X}(\sin^2\Theta)\mathbf{X}^\top.\end{aligned}\tag{2.36}$$

Here, the penultimate identity follows from our construction (cf. (2.5)), where we define $\cos\Theta := \text{diag}([\cos\theta_1, \dots, \cos\theta_r])$. Therefore, for any $1 \leq i \leq r$, the i -th largest singular value of $\mathbf{U}^\top\mathbf{U}_\perp^*$ obeys

$$\sigma_i(\mathbf{U}^\top\mathbf{U}_\perp^*) = \sqrt{\lambda_i(\mathbf{U}^\top\mathbf{U}_\perp^*\mathbf{U}_\perp^*\top\mathbf{U})} = \sin\theta_{r+1-i},$$

which results from the ordering in (2.6). This means that, if $r \leq n - r$, then the singular values of $\mathbf{U}^\top\mathbf{U}_\perp^*$ are precisely given by $\{\sin\theta_i\}_{1 \leq i \leq r}$. Repeating this argument reveals that the singular values of $\mathbf{U}_\perp^*\mathbf{U}^*$ are also $\{\sin\theta_i\}_{1 \leq i \leq r}$ if $r \leq n - r$.

Combining the above observations thus completes the proof.

2.6.2 Proof of Lemma 2.5

A closer inspection of the proof of Lemma 2.4 (in particular, (2.36) and the orthonormality of \mathbf{X}) reveals that

$$\begin{aligned}\|\mathbf{U}^\top\mathbf{U}_\perp^*\| &= \sqrt{\|\mathbf{U}^\top\mathbf{U}_\perp^*\mathbf{U}_\perp^*\top\mathbf{U}\|} = \sqrt{\|\mathbf{X}(\sin^2\Theta)\mathbf{X}^\top\|} = \|\sin\Theta\|, \\ \|\mathbf{U}^\top\mathbf{U}_\perp^*\|_F &= \sqrt{\text{Tr}(\mathbf{U}^\top\mathbf{U}_\perp^*\mathbf{U}_\perp^*\top\mathbf{U})} = \sqrt{\text{Tr}(\mathbf{X}(\sin^2\Theta)\mathbf{X}^\top)} \\ &= \sqrt{\text{Tr}(\mathbf{X}^\top\mathbf{X}\sin^2\Theta)} = \sqrt{\text{Tr}(\sin^2\Theta)} = \|\sin\Theta\|_F,\end{aligned}$$

where we have used the basic property $\text{Tr}(\mathbf{AB}) = \text{Tr}(\mathbf{BA})$. Similarly,

$$\|\mathbf{U}_\perp^*\mathbf{U}^*\| = \|\sin\Theta\|, \quad \text{and} \quad \|\mathbf{U}_\perp^*\mathbf{U}^*\|_F = \|\sin\Theta\|_F.$$

Note that the above identities hold for all $1 \leq r \leq n$. In addition, the relation (2.35) tells us that

$$\|\mathbf{U}\mathbf{U}^\top - \mathbf{U}^*\mathbf{U}^{*\top}\| = \max\{\|\mathbf{U}^\top\mathbf{U}_\perp^*\|, \|\mathbf{U}_\perp^\top\mathbf{U}^*\|\}; \quad (2.37a)$$

$$\|\mathbf{U}\mathbf{U}^\top - \mathbf{U}^*\mathbf{U}^{*\top}\|_F = \left(\|\mathbf{U}^\top\mathbf{U}_\perp^*\|_F^2 + \|\mathbf{U}_\perp^\top\mathbf{U}^*\|_F^2\right)^{1/2}. \quad (2.37b)$$

Putting the above identities together immediately establishes the advertised results.

2.6.3 Proof of Lemma 2.6

As before, suppose that the SVD of $\mathbf{U}^\top\mathbf{U}^*$ is given by $\mathbf{X}\mathbf{\Sigma}\mathbf{Y}^\top$, where \mathbf{X} and \mathbf{Y} are $r \times r$ orthonormal matrices whose columns contain the left singular vectors and the right singular vectors of $\mathbf{U}^\top\mathbf{U}^*$, respectively, and $\mathbf{\Sigma} \in \mathbb{R}^{r \times r} = \cos \Theta$ is a diagonal matrix whose diagonal entries correspond to the singular values of $\mathbf{U}^\top\mathbf{U}^*$.

The spectral norm upper bound. We first observe that

$$\begin{aligned} \|\mathbf{U}\mathbf{X}\mathbf{Y}^\top - \mathbf{U}^*\|^2 &= \|(\mathbf{U}\mathbf{X}\mathbf{Y}^\top - \mathbf{U}^*)^\top(\mathbf{U}\mathbf{X}\mathbf{Y}^\top - \mathbf{U}^*)\| \\ &= \|2\mathbf{I}_r - \mathbf{Y}\mathbf{X}^\top\mathbf{U}^\top\mathbf{U}^* - \mathbf{U}^{*\top}\mathbf{U}\mathbf{X}\mathbf{Y}^\top\| \\ &= \|2\mathbf{I}_r - \mathbf{Y}\mathbf{X}^\top\mathbf{X}\mathbf{\Sigma}\mathbf{Y}^\top - \mathbf{Y}\mathbf{\Sigma}\mathbf{X}^\top\mathbf{X}\mathbf{Y}^\top\| \\ &= 2\|\mathbf{Y}(\mathbf{I}_r - \mathbf{\Sigma})\mathbf{Y}^\top\| = 2\|\mathbf{I}_r - \mathbf{\Sigma}\|. \end{aligned} \quad (2.38)$$

Here, the penultimate line relies on the singular value decomposition $\mathbf{U}^\top\mathbf{U}^* = \mathbf{X}\mathbf{\Sigma}\mathbf{Y}^\top$, while the two identities in the last line result from the orthonormality of \mathbf{X} and \mathbf{Y} , respectively. In addition, note that

$$\begin{aligned} \|\mathbf{I}_r - \mathbf{\Sigma}\| &= \|\mathbf{I}_r - \cos \Theta\| \leq \|\mathbf{I}_r - \cos^2 \Theta\| \\ &= \|\sin^2 \Theta\| = \|\sin \Theta\|^2. \end{aligned}$$

This taken together with (2.38) leads to

$$\min_{\mathbf{R} \in \mathcal{O}^{r \times r}} \|\mathbf{U}\mathbf{R} - \mathbf{U}^*\| \leq \|\mathbf{U}\mathbf{X}\mathbf{Y}^\top - \mathbf{U}^*\| \leq \sqrt{2}\|\sin \Theta\|,$$

where the first inequality holds since \mathbf{X} and \mathbf{Y} are both orthonormal matrices and hence $\mathbf{X}\mathbf{Y}^\top$ is also orthonormal.

The spectral norm lower bound. On the other hand, we make the observation that

$$\begin{aligned} \min_{\mathbf{R} \in \mathcal{O}^{r \times r}} \|\mathbf{U}\mathbf{R} - \mathbf{U}^*\|^2 &= \min_{\mathbf{R} \in \mathcal{O}^{r \times r}} \|(\mathbf{U}\mathbf{R} - \mathbf{U}^*)^\top(\mathbf{U}\mathbf{R} - \mathbf{U}^*)\| \\ &= \min_{\mathbf{R} \in \mathcal{O}^{r \times r}} \|\mathbf{R}^\top \mathbf{U}^\top \mathbf{U}\mathbf{R} + \mathbf{U}^{*\top} \mathbf{U}^* - \mathbf{R}^\top \mathbf{U}^\top \mathbf{U}^* - \mathbf{U}^{*\top} \mathbf{U}\mathbf{R}\| \\ &= \min_{\mathbf{R} \in \mathcal{O}^{r \times r}} \|2\mathbf{I}_r - \mathbf{R}^\top \mathbf{X}\Sigma\mathbf{Y}^\top - \mathbf{Y}\Sigma\mathbf{X}^\top \mathbf{R}\|, \end{aligned} \quad (2.39)$$

where the last relation holds since $\mathbf{X}\Sigma\mathbf{Y}^\top$ is the SVD of $\mathbf{U}^\top \mathbf{U}^*$. Continue the derivation to obtain

$$\begin{aligned} (2.39) &\stackrel{(i)}{=} \min_{\mathbf{Q} \in \mathcal{O}^{r \times r}} \|2\mathbf{I}_r - \mathbf{Q}\Sigma\mathbf{Y}^\top - \mathbf{Y}\Sigma\mathbf{Q}^\top\| \\ &\stackrel{(ii)}{=} \min_{\mathbf{Q} \in \mathcal{O}^{r \times r}} \|2\mathbf{Q}^\top \mathbf{Q} - \mathbf{Q}^\top \mathbf{Q}\Sigma\mathbf{Y}^\top \mathbf{Q} - \mathbf{Q}^\top \mathbf{Y}\Sigma\mathbf{Q}^\top \mathbf{Q}\| \\ &= \min_{\mathbf{Q} \in \mathcal{O}^{r \times r}} \|2\mathbf{I}_r - \Sigma\mathbf{Y}^\top \mathbf{Q} - \mathbf{Q}^\top \mathbf{Y}\Sigma\| \\ &\stackrel{(iii)}{=} \min_{\mathbf{O} \in \mathcal{O}^{r \times r}} \|2\mathbf{I}_r - \Sigma\mathbf{O} - \mathbf{O}^\top \Sigma\|. \end{aligned} \quad (2.40)$$

Here, (i) follows by setting $\mathbf{Q} = \mathbf{R}^\top \mathbf{X}$ (since both \mathbf{X} and \mathbf{R} are orthonormal matrices), (ii) results from the unitary invariance of the spectral norm, whereas (iii) holds by setting $\mathbf{O} = \mathbf{Y}^\top \mathbf{Q}$. Moreover, recognizing that $\|\Sigma\mathbf{O}\| \leq \|\Sigma\| \cdot \|\mathbf{O}\| \leq 1$ (and hence $2\mathbf{I}_r - \Sigma\mathbf{O} - \mathbf{O}^\top \Sigma \succeq \mathbf{0}$), one can obtain

$$\begin{aligned} \min_{\mathbf{O} \in \mathcal{O}^{r \times r}} \|2\mathbf{I}_r - \Sigma\mathbf{O} - \mathbf{O}^\top \Sigma\| &= \min_{\mathbf{O} \in \mathcal{O}^{r \times r}} \lambda_{\max}(2\mathbf{I}_r - \Sigma\mathbf{O} - \mathbf{O}^\top \Sigma) \\ &= \min_{\mathbf{O} \in \mathcal{O}^{r \times r}} \max_{\mathbf{u}: \|\mathbf{u}\|_2=1} \mathbf{u}^\top (2\mathbf{I}_r - \Sigma\mathbf{O} - \mathbf{O}^\top \Sigma) \mathbf{u} \\ &= \min_{\mathbf{O} \in \mathcal{O}^{r \times r}} \max_{\mathbf{u}: \|\mathbf{u}\|_2=1} (2 - 2\mathbf{u}^\top \Sigma\mathbf{O}\mathbf{u}) \\ &\geq \min_{\mathbf{O} \in \mathcal{O}^{r \times r}} (2 - 2\mathbf{e}_r^\top \Sigma\mathbf{O}\mathbf{e}_r) \\ &= 2 - 2 \cos \theta_r \max_{\mathbf{O} \in \mathcal{O}^{r \times r}} \mathbf{e}_r^\top \mathbf{O}\mathbf{e}_r \\ &\geq 2 - 2 \cos \theta_r = 4 \sin^2(\theta_r/2). \end{aligned} \quad (2.41)$$

Here, the inequality follows by taking \mathbf{u} to be \mathbf{e}_r (recall that by construction, $\sigma_r = \cos \theta_r \geq 0$ is the smallest singular value of Σ), and the

penultimate line holds by combining the facts $|e_r^\top \mathbf{O} e_r| \leq \|\mathbf{O}\| = 1$ and $e_r^\top e_r = 1$. Putting (2.41) and (2.40) together yields

$$\begin{aligned} \min_{\mathbf{R} \in \mathcal{O}^{r \times r}} \|\mathbf{U}\mathbf{R} - \mathbf{U}^*\| &\geq \sqrt{4 \sin^2(\theta_r/2)} = 2 \sin(\theta_r/2) = \|2 \sin(\Theta/2)\| \\ &\geq \|\sin \Theta\|, \end{aligned}$$

where we again use the inequality $2 \sin(\theta/2) \geq \sin \theta$ for all $\theta \in [0, \pi/2]$.

Finally, invoking the relation $\|\sin \Theta\| = \|\mathbf{U}\mathbf{U}^\top - \mathbf{U}^*\mathbf{U}^{*\top}\|$ (see Lemma 2.5) establishes the claimed spectral norm bounds.

The Frobenius norm upper bound. Regarding the Frobenius norm upper bound, one sees that

$$\begin{aligned} \|\mathbf{U}\mathbf{X}\mathbf{Y}^\top - \mathbf{U}^*\|_{\text{F}}^2 &= \|\mathbf{U}\|_{\text{F}}^2 + \|\mathbf{U}^*\|_{\text{F}}^2 - 2\text{Tr}(\mathbf{Y}\mathbf{X}^\top \mathbf{U}^\top \mathbf{U}^*) \\ &\stackrel{(i)}{=} r + r - 2\text{Tr}(\mathbf{Y}\mathbf{X}^\top \mathbf{X}\Sigma\mathbf{Y}^\top) \stackrel{(ii)}{=} 2r - 2\text{Tr}(\Sigma), \end{aligned} \quad (2.42)$$

where (i) holds since \mathbf{U} and \mathbf{U}^* are both $n \times r$ matrices with orthonormal columns, and (ii) follows since $\mathbf{X}^\top \mathbf{X} = \mathbf{Y}^\top \mathbf{Y} = \mathbf{I}$ (and hence $\text{Tr}(\mathbf{Y}\mathbf{X}^\top \mathbf{X}\Sigma\mathbf{Y}^\top) = \text{Tr}(\mathbf{Y}^\top \mathbf{Y}\mathbf{X}^\top \mathbf{X}\Sigma) = \text{Tr}(\Sigma)$). Furthermore,

$$\begin{aligned} 2r - 2\text{Tr}(\Sigma) &\stackrel{(iii)}{=} 2 \sum_i (1 - \cos \theta_i) \leq 2 \sum_i (1 - \cos^2 \theta_i) \\ &= 2 \|\sin \Theta\|_{\text{F}}^2 = \|\mathbf{U}\mathbf{U}^\top - \mathbf{U}^*\mathbf{U}^{*\top}\|_{\text{F}}^2, \end{aligned}$$

where (iii) holds by construction (cf. (2.5)), and the last identity results from Lemma 2.5. This taken collectively with (2.42) reveals that

$$\min_{\mathbf{R} \in \mathcal{O}^{r \times r}} \|\mathbf{U}\mathbf{R} - \mathbf{U}^*\|_{\text{F}}^2 \leq \|\mathbf{U}\mathbf{X}\mathbf{Y}^\top - \mathbf{U}^*\|_{\text{F}}^2 \leq \|\mathbf{U}\mathbf{U}^\top - \mathbf{U}^*\mathbf{U}^{*\top}\|_{\text{F}}^2,$$

where the first inequality holds since \mathbf{X} and \mathbf{Y} are both orthonormal matrices and hence $\mathbf{X}\mathbf{Y}^\top$ is also orthonormal.

The Frobenius norm lower bound. With regards to the Frobenius norm lower bound, it is seen that

$$\begin{aligned} \min_{\mathbf{R} \in \mathcal{O}^{r \times r}} \|\mathbf{U}\mathbf{R} - \mathbf{U}^*\|_{\text{F}}^2 &= \min_{\mathbf{R} \in \mathcal{O}^{r \times r}} \left\{ \|\mathbf{U}\mathbf{R}\|_{\text{F}}^2 + \|\mathbf{U}^*\|_{\text{F}}^2 - 2\langle \mathbf{U}\mathbf{R}, \mathbf{U}^* \rangle \right\} \\ &\stackrel{(i)}{=} 2 \min_{\mathbf{R} \in \mathcal{O}^{r \times r}} \left\{ r - \langle \mathbf{R}, \mathbf{U}^\top \mathbf{U}^* \rangle \right\} \end{aligned}$$

$$\stackrel{(ii)}{=} 2 \min_{\mathbf{R} \in \mathcal{O}^{r \times r}} \left\{ r - \langle \mathbf{R}, \mathbf{X} \Sigma \mathbf{Y}^\top \rangle \right\}, \quad (2.43)$$

where (i) holds since $\|\mathbf{U}\|_{\text{F}} = \|\mathbf{U}^*\|_{\text{F}} = \sqrt{r}$, and (ii) relies on the SVD $\mathbf{X} \Sigma \mathbf{Y}^\top$ of $\mathbf{U}^\top \mathbf{U}^*$. Continue the derivation to obtain

$$\begin{aligned} (2.43) &\stackrel{(iii)}{=} 2 \min_{\mathbf{Q} \in \mathcal{O}^{r \times r}} \left\{ r - \langle \mathbf{Q}, \cos \Theta \rangle \right\} \stackrel{(iv)}{\geq} 2 \min_{\mathbf{Q} \in \mathcal{O}^{r \times r}} \left\{ r - \|\mathbf{Q}\| \|\cos \Theta\|_* \right\} \\ &= 2(r - \sum_i \cos \theta_i). \end{aligned} \quad (2.44)$$

Here, (iii) sets $\mathbf{Q} = \mathbf{X}^\top \mathbf{R} \mathbf{Y}$ and identifies Σ as $\cos \Theta$, (iv) comes from the elementary inequality $\langle \mathbf{A}, \mathbf{B} \rangle \leq \|\mathbf{A}\| \|\mathbf{B}\|_*$, whereas the last line follows since $\cos \theta_i \geq 0$. Additionally, it is easily seen that

$$\begin{aligned} (2.44) &= 2 \sum_i (1 - \cos \theta_i) = 4 \sum_i \sin^2(\theta_i/2) \\ &\geq \sum_i \sin^2 \theta_i = \frac{1}{2} \|\mathbf{U} \mathbf{U}^\top - \mathbf{U}^* \mathbf{U}^{*\top}\|_{\text{F}}^2, \end{aligned} \quad (2.45)$$

where the penultimate relation follows from the elementary inequality $2 \sin(\theta/2) \geq \sin \theta$ (which holds for any $0 \leq \theta \leq \pi/2$), and the last line invokes Lemma 2.5. Combining the inequalities (2.44) and (2.45), we establish the claimed lower bound.

2.7 Notes

Additional resources on matrix perturbation theory. Matrix perturbation theory is a firmly established topic that has been extensively studied in the past several decades. Two classic books that offer in-depth discussions of perturbation theory for eigenspaces and singular subspaces are Stewart and Sun (1990) and Sun (1987). Other valuable resources on this topic include Bhatia (2013) and Horn and Johnson (2012). The exposition herein is largely influenced by the excellent lecture notes by Montanari (2011) and Hsu (2016). In addition, the book (Kato, 2013) offers a more abstract treatment of perturbation theory from the viewpoint of linear operators. Several variants of the $\sin \Theta$ theorem amenable to statistical analysis are available in the statistics literature as well (e.g., Yu *et al.* (2015), Vu and Lei (2013), Cai and Zhang (2018), and Zhang *et al.* (2021)).

Extensions. We point out several well-known extensions of the theorems presented in this chapter. To begin with, the current exposition restricts attention to the real case for simplicity, while in fact all results herein generalize to the complex-valued case (Stewart and Sun, 1990). In addition, Theorem 2.9 together with Lemma 2.6 reveals the existence of two rotation matrices \mathbf{R}_U and \mathbf{R}_V obeying

$$\max \{ \|\mathbf{U}\mathbf{R}_U - \mathbf{U}^*\|_{\text{F}}, \|\mathbf{V}\mathbf{R}_V - \mathbf{V}^*\|_{\text{F}} \} \leq \frac{\sqrt{2} \max \{ \|\mathbf{E}^\top \mathbf{U}^*\|_{\text{F}}, \|\mathbf{E}\mathbf{V}^*\|_{\text{F}} \}}{\sigma_r^* - \sigma_{r+1}^* - \|\mathbf{E}\|},$$

but falls short of illuminating the connection between \mathbf{R}_U and \mathbf{R}_V . An extension derived in Dopico (2000) establishes a similar perturbation bound even when \mathbf{R}_U and \mathbf{R}_V are taken to be the same rotation matrix.

3

Applications of ℓ_2 perturbation theory to data science

This chapter develops tailored spectral methods for several important applications arising in statistics, machine learning and signal processing. As it turns out, these methods are all variations of a common recipe: extracting the information of interest from the eigenspace (resp. singular spaces) and eigenvalues (resp. singular values) of a certain matrix \mathbf{M} properly constructed from data. The inspiration stems from the observation that: the corresponding quantities of $\mathbf{M}^* = \mathbb{E}[\mathbf{M}]$ —when properly constructed and under appropriate statistical models—might faithfully reveal the information being sought after. The classical ℓ_2 perturbation theory introduced in Chapter 2, when paired with modern probabilistic tools reviewed in Section 3.1, uncovers appealing performance of spectral methods in numerous applications by controlling the size of the perturbation $\mathbf{E} := \mathbf{M} - \mathbf{M}^*$. The vignettes in this chapter provide ample evidence regarding the benefits of harnessing the statistical nature of the acquired data.

3.1 Preliminaries: Matrix tail bounds

In order to invoke the $\sin\Theta$ theorems (Theorems 2.7 and 2.9), an important ingredient lies in developing a tight upper bound on the

spectral norm $\|\mathbf{E}\|$ of the perturbation matrix \mathbf{E} . This is where statistical/probabilistic tools play a major role. Rather than presenting an encyclopedia of probabilistic techniques (which can be gleaned from Tropp (2015), Vershynin (2017), Boucheron *et al.* (2013), Wainwright (2019), Tropp (2011), Raginsky and Sason (2013), and Howard *et al.* (2020)), this monograph singles out only two useful matrix concentration inequalities that suffice for the applications considered herein.

The (truncated) matrix Bernstein inequality

The first result is an extension of the celebrated matrix Bernstein inequality (Oliveira, 2009; Tropp, 2012; Hopkins *et al.*, 2016). This is an elegant and convenient tail bound for the sum of independent random matrices, resulting in effective performance guarantees for a diverse array of statistical applications. We refer the interested reader to Tropp (2015) for a highly accessible introduction of the classical matrix Bernstein inequality, and Hopkins *et al.* (2016, Section A.2.2) for a proof of the truncated variant stated in Theorem 3.1.

Theorem 3.1 (Truncated matrix Bernstein). Let $\{\mathbf{X}_i\}_{1 \leq i \leq m}$ be a sequence of independent real random matrices with dimension $n_1 \times n_2$. Suppose that for all $1 \leq i \leq m$,

$$\mathbb{P}\{\|\mathbf{X}_i - \mathbb{E}[\mathbf{X}_i]\| \geq L\} \leq q_0 \quad (3.1a)$$

$$\|\mathbb{E}[\mathbf{X}_i] - \mathbb{E}[\mathbf{X}_i \mathbf{1}\{\|\mathbf{X}_i\| < L\}]\| \leq q_1 \quad (3.1b)$$

hold for some quantities $0 \leq q_0 \leq 1$ and $q_1 \geq 0$. In addition, define the matrix variance statistic v as

$$v := \max \left\{ \left\| \sum_{i=1}^m \mathbb{E}[(\mathbf{X}_i - \mathbb{E}[\mathbf{X}_i])(\mathbf{X}_i - \mathbb{E}[\mathbf{X}_i])^\top] \right\|, \right. \\ \left. \left\| \sum_{i=1}^m \mathbb{E}[(\mathbf{X}_i - \mathbb{E}[\mathbf{X}_i])^\top(\mathbf{X}_i - \mathbb{E}[\mathbf{X}_i])] \right\| \right\}. \quad (3.2)$$

Then for all $t \geq mq_1$, one has

$$\mathbb{P}\left(\left\|\sum_{i=1}^m (\mathbf{X}_i - \mathbb{E}[\mathbf{X}_i])\right\| \geq t\right) \leq (n_1 + n_2) \exp\left(\frac{-(t - mq_1)^2/2}{v + L(t - mq_1)/3}\right) + mq_0.$$

Remark 3.1. Note that when the \mathbf{X}_i 's are i.i.d. zero-mean random matrices, the matrix variance statistic simplifies to

$$v = m \max \left\{ \|\mathbb{E}[\mathbf{X}_i \mathbf{X}_i^\top]\|, \|\mathbb{E}[\mathbf{X}_i^\top \mathbf{X}_i]\| \right\}.$$

To make it more user-friendly, we record a straightforward consequence of Theorem 3.1 as follows.

Corollary 3.2. Suppose the assumptions of Theorem 3.1 hold, and set $n := \max\{n_1, n_2\}$. For any $a \geq 2$, with probability exceeding $1 - 2n^{-a+1} - mq_0$ one has

$$\left\| \sum_{i=1}^m (\mathbf{X}_i - \mathbb{E}[\mathbf{X}_i]) \right\| \leq \sqrt{2av \log n} + \frac{2a}{3} L \log n + mq_1. \quad (3.3)$$

In order to make effective use of the above results (particularly when handling unbounded random matrices), it is advisable to take L as a high-probability bound on $\|\mathbf{X}_i - \mathbb{E}[\mathbf{X}_i]\|$. The rationale is simple: by properly truncating \mathbf{X}_i based on the level L , we end up with a *bounded* sequence that is more convenient to work with while not deviating much from the original sequence. In particular, if all $\|\mathbf{X}_i - \mathbb{E}[\mathbf{X}_i]\|$ are bounded by a deterministic quantity which is set to be L , then both q_0 and q_1 vanish, thus eliminating the need of enforcing truncation. In this case, Corollary 3.2 simplifies to a user-friendly version of the standard matrix Bernstein inequality, which we record below for ease of reference.

Corollary 3.3 (Matrix Bernstein). Let $\{\mathbf{X}_i\}_{1 \leq i \leq m}$ be a set of independent real random matrices with dimension $n_1 \times n_2$. Suppose that

$$\mathbb{E}[\mathbf{X}_i] = \mathbf{0}, \quad \text{and} \quad \|\mathbf{X}_i\| \leq L, \quad \text{for all } i. \quad (3.4)$$

Set $n := \max\{n_1, n_2\}$, and recall the definition of variance statistic in (3.2). For any $a \geq 2$, with probability exceeding $1 - 2n^{-a+1}$ one has

$$\left\| \sum_{i=1}^m \mathbf{X}_i \right\| \leq \sqrt{2av \log n} + \frac{2a}{3} L \log n. \quad (3.5)$$

By virtue of the above inequalities, the key to bounding $\|\sum_i \mathbf{X}_i\|$ largely lies in controlling the following two crucial quantities:

$$\sqrt{v \log n} \quad \text{and} \quad L \log n,$$

where the former depends on the number m of random matrices involved.

Spectral norm of random matrices with independent entries

An important family of random matrices that merits special attention comprises the ones with independent random entries, that is, matrices of the form $\mathbf{X} = [X_{i,j}]_{1 \leq i,j \leq n}$ with independent $X_{i,j}$'s. While the spectral norm of such a matrix can also be analyzed via matrix Bernstein (by treating \mathbf{X} as the sum of independent random matrices $X_{i,j} \mathbf{e}_i \mathbf{e}_j^\top$), this approach is typically loose in terms of the logarithmic factor. Motivated by the abundance of such random matrices in practice, we record below a strengthened non-asymptotic spectral norm bound, which is of significant utility and is tighter than what matrix Bernstein has to offer for this case.

Theorem 3.4. Consider a *symmetric* random matrix $\mathbf{X} = [X_{i,j}]_{1 \leq i,j \leq n}$ in $\mathbb{R}^{n \times n}$, whose entries are independently generated and obey

$$\mathbb{E}[X_{i,j}] = 0, \quad \text{and} \quad |X_{i,j}| \leq B, \quad 1 \leq i, j \leq n. \quad (3.6)$$

Define

$$\nu := \max_i \sum_j \mathbb{E}[X_{i,j}^2]. \quad (3.7)$$

Then there exists some universal constant $c > 0$ such that for any $t \geq 0$,

$$\mathbb{P}\left\{ \|\mathbf{X}\| \geq 4\sqrt{\nu} + t \right\} \leq n \exp\left(-\frac{t^2}{cB^2}\right). \quad (3.8)$$

This result, which appeared in Bandeira and Van Handel (2016, Remark 3.13), can be established via tighter control of the expected spectral norm in conjunction with Talagrand's concentration inequality. Two remarks are in order.

- First, it is easy to see that the result extends to asymmetric matrices with independent entries, using the standard “dilation trick” (see, e.g., (Tropp, 2015, Section 2.1.17)). Specifically, for an asymmetric random matrix $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2}$, let us introduce the symmetric dilation $\mathcal{S}(\mathbf{X})$ of \mathbf{X} :

$$\mathcal{S}(\mathbf{X}) := \begin{bmatrix} \mathbf{0} & \mathbf{X} \\ \mathbf{X}^\top & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{(n_1+n_2) \times (n_1+n_2)},$$

which enjoys the desired symmetry and can be analyzed directly using Theorem 3.4. The resulting bound on $\|\mathcal{S}(\mathbf{X})\|$ can be translated back to $\|\mathbf{X}\|$ via the elementary identity $\|\mathbf{X}\| = \|\mathcal{S}(\mathbf{X})\|$. For conciseness, we will occasionally apply Theorem 3.4 directly to asymmetric matrices without invoking the dilation trick.

- As a useful corollary, if we know *a priori* that $\mathbb{E}[X_{i,j}^2] \leq \sigma^2$ for all $1 \leq i, j \leq n$, then Theorem 3.4 implies that

$$\|\mathbf{X}\| \leq 4\sigma\sqrt{n} + \tilde{c}B\sqrt{\log n} \quad (3.9)$$

with probability at least $1 - n^{-8}$ for some constant $\tilde{c} > 0$. To see this, it suffices to set $\tilde{c} = \sqrt{9c}$ and take $t = B\sqrt{9c\log n}$ in (3.8).

Remark 3.2. The inequality (3.9) continues to hold if we replace n^{-8} with $n^{-\alpha}$ for any positive constant $\alpha > 0$. Here and below, we often go with the artificial choice like n^{-8} since it is small enough for our purpose.

3.2 Low-rank matrix denoising

To catch a glimpse of the effectiveness of the approach we have introduced, let us start by trying it out on a warm-up example: low-rank matrix denoising.

3.2.1 Problem formulation and algorithm

Consider an unknown rank- r symmetric matrix $\mathbf{M}^* \in \mathbb{R}^{n \times n}$ with eigendecomposition $\mathbf{M}^* = \mathbf{U}^* \boldsymbol{\Lambda}^* \mathbf{U}^{*\top}$, where the columns of $\mathbf{U}^* \in \mathbb{R}^{n \times r}$ are orthonormal, and $\boldsymbol{\Lambda}^* \in \mathbb{R}^{r \times r}$ is a diagonal matrix containing the nonzero eigenvalues $\{\lambda_i^*\}$ of \mathbf{M}^* . Assume that $|\lambda_1^*| \geq |\lambda_2^*| \geq \dots \geq |\lambda_r^*| > 0$. Suppose that we observe a noisy copy

$$\mathbf{M} = \mathbf{M}^* + \mathbf{E},$$

where $\mathbf{E} = [E_{i,j}]_{1 \leq i,j \leq n}$ is a symmetric noise matrix. It is assumed that the entries $\{E_{i,j}\}_{i \geq j}$ are independently generated obeying

$$E_{i,j} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2), \quad i \geq j. \quad (3.10)$$

The aim is to estimate the eigenspace \mathbf{U}^* from the data matrix \mathbf{M} . Despite its simplicity, this problem has been extensively studied in the literature (Koltchinskii and Xia, 2016; Bao *et al.*, 2021; Ding, 2020; Xia, 2021; Li *et al.*, 2021a). It also bears close relevance to the famous angular/phase synchronization problem (Singer, 2011; Bandeira *et al.*, 2017).

In order to estimate the low-rank factors specified by \mathbf{U}^* , a natural scheme is to resort to the rank- r leading eigenspace of the data matrix \mathbf{M} . More precisely, denote by $\lambda_1, \dots, \lambda_n$ the eigenvalues of \mathbf{M} sorted by their magnitudes, i.e.,

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|, \quad (3.11)$$

and let $\mathbf{u}_1, \dots, \mathbf{u}_n$ represent the associated eigenvectors. This spectral method returns $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_r] \in \mathbb{R}^{n \times r}$ as an estimate of \mathbf{U}^* .

3.2.2 Performance guarantees

Statistical accuracy of the spectral estimate. We now examine the accuracy of the above spectral estimate. Towards this, a key step lies in bounding the spectral norm of the noise matrix \mathbf{E} . We claim for the moment that (which will be established in Section 3.2.3)

$$\|\mathbf{E}\| \leq 5\sigma\sqrt{n} \quad (3.12)$$

with probability at least $1 - O(n^{-8})$. Armed with this claim and the fact $\lambda_{r+1}^* = 0$, we are in a situation where it is quite easy to see how the Davis-Kahan theorem applies. According to Corollary 2.8, with probability greater than $1 - O(n^{-8})$ one has

$$\text{dist}(\mathbf{U}, \mathbf{U}^*) \leq \frac{2\|\mathbf{E}\|}{|\lambda_r^*|} \leq \frac{10\sigma\sqrt{n}}{|\lambda_r^*|}, \quad (3.13)$$

provided that the noise variance is sufficiently small obeying $\sigma\sqrt{n} \leq \frac{1-1/\sqrt{2}}{5}|\lambda_r^*|$ so that $\|\mathbf{E}\| \leq (1 - 1/\sqrt{2})|\lambda_r^*|$.

Tightness and optimality. The tightness of the statistical guarantee (3.13) can be assessed when compared with the minimax lower bound. For instance, it is well-known in the literature (e.g., Cheng *et al.*

(2021, Theorem 3)) that: even for the case with $r = 1$, one cannot hope to achieve $\text{dist}(\widehat{\mathbf{U}}, \mathbf{U}^*) = o(\sigma\sqrt{n}/|\lambda_r^*|)$ —in a minimax sense—regardless of the estimator $\widehat{\mathbf{U}}$ in use. Consequently, the spectral method turns out to be orderwise statistically optimal for low-rank matrix denoising.

Additional useful results: eigenvalue and matrix estimation. Before concluding, we record several immediate consequences of the above analysis that will be useful later on. Specifically, assuming that $\sigma\sqrt{n} \leq \frac{1-1/\sqrt{2}}{5}|\lambda_r^*|$, we see from Weyl’s inequality (cf. Lemma 2.2) that

$$|\lambda_i| \leq \|\mathbf{E}\| \leq 5\sigma\sqrt{n}, \quad \text{for all } i \geq r+1 \quad (3.14)$$

with probability $1 - O(n^{-8})$.

We further remark on the Euclidean statistical accuracy when estimating the unknown matrix \mathbf{M}^* using $\widehat{\mathbf{M}} := \mathbf{U}\Lambda\mathbf{U}^\top$, where $\Lambda := \text{diag}([\lambda_1, \dots, \lambda_r])$. It is seen from the triangle inequality that

$$\begin{aligned} \|\mathbf{U}\Lambda\mathbf{U}^\top - \mathbf{M}^*\| &\leq \|\mathbf{M} - \mathbf{M}^*\| + \|\mathbf{U}\Lambda\mathbf{U}^\top - \mathbf{M}\| \\ &= \|\mathbf{E}\| + |\lambda_{r+1}| \leq 2\|\mathbf{E}\|, \end{aligned} \quad (3.15)$$

where the last inequality relies on (3.14). Since the rank of $\mathbf{U}\Lambda\mathbf{U}^\top - \mathbf{M}^*$ is at most $2r$, with probability at least $1 - O(n^{-8})$ one has

$$\begin{aligned} \|\mathbf{U}\Lambda\mathbf{U}^\top - \mathbf{M}^*\|_{\text{F}} &\leq \sqrt{2r} \|\mathbf{U}\Lambda\mathbf{U}^\top - \mathbf{M}^*\| \leq 2\sqrt{2r} \|\mathbf{E}\| \\ &\leq 10\sigma\sqrt{2nr}. \end{aligned} \quad (3.16)$$

3.2.3 Proof of the inequality (3.12) on $\|\mathbf{E}\|$

We plan to employ Theorem 3.4. Given that Gaussian entries are unbounded, we introduce a truncated copy $\tilde{\mathbf{E}} = [\tilde{E}_{i,j}]_{1 \leq i,j \leq n}$ defined as follows

$$\tilde{E}_{i,j} := E_{i,j} \mathbb{1}\{|E_{i,j}| \leq 5\sigma\sqrt{\log n}\}, \quad 1 \leq i, j \leq n. \quad (3.17)$$

Two properties are in place.

- It is readily seen from the property of Gaussian distributions that

$$\mathbb{P}\{E_{i,j} = \tilde{E}_{i,j}\} \geq 1 - n^{-12}, \quad 1 \leq i, j \leq n,$$

which combined with the union bound leads to

$$\mathbb{P}\{\mathbf{E} = \tilde{\mathbf{E}}\} \geq 1 - n^{-10}. \quad (3.18)$$

- Given that $B := \max_{i,j} |\tilde{E}_{i,j}| \leq 5\sigma\sqrt{\log n}$, we can invoke Theorem 3.4 (or more directly, (3.9)) to demonstrate that

$$\|\tilde{\mathbf{E}}\| \leq 4\sigma\sqrt{n} + O(B\log n) \leq 5\sigma\sqrt{n}$$

for sufficiently large n , with probability exceeding $1 - O(n^{-8})$. Here, we implicitly use the fact that $\mathbb{E}[\tilde{E}_{i,j}^2] \leq \mathbb{E}[E_{i,j}^2] = \sigma^2$.

Combining the above two observations implies that

$$\|\mathbf{E}\| = \|\tilde{\mathbf{E}}\| \leq 5\sigma\sqrt{n}$$

with probability exceeding $1 - O(n^{-8})$, as claimed.

3.3 Principal component analysis and factor models

Principal component analysis (PCA) and factor models (Jolliffe, 1986; Lawley and Maxwell, 1962; Fan *et al.*, 2020b)—which serve as an effective unsupervised learning tool for exploring and understanding data—arise frequently in data-intensive applications in economics, finance, psychology, signal processing, speech, neuroscience, traffic data analysis, among other things (Stock and Watson, 2002; McCrae and John, 1992; Scharf, 1991; Chen *et al.*, 2015a; Balzano *et al.*, 2018; Fan *et al.*, 2021b). PCA and factor models not only allow for dimensionality reduction, but also provide intermediate means for data visualization, noise removal, anomaly detection, and other downstream tasks. In this section, we investigate a simple, yet broadly applicable, factor model.

3.3.1 Problem formulation and assumptions

Dependence of high-dimensional measurements is a stylized feature in data science. To model the dependence among observed high-dimensional data, we assume that there are latent factors that drive the dependence, with a loading matrix that describes how each component depends on the latent factors and an idiosyncratic noise that captures the remaining

part. To set the stage, imagine we have collected a set of n independent sample vectors $\mathbf{x}_i \in \mathbb{R}^p$, $1 \leq i \leq n$ obeying

$$\mathbf{x}_i = \mathbf{L}^* \mathbf{f}_i + \boldsymbol{\eta}_i, \quad 1 \leq i \leq n. \quad (3.19)$$

Here, $\mathbf{f}_i \in \mathbb{R}^r$ is a vector of latent factors, $\mathbf{L}^* \in \mathbb{R}^{p \times r}$ represents a factor loading matrix that is not known *a priori*, whereas $\boldsymbol{\eta}_i \in \mathbb{R}^p$ stands for additive random noise or the idiosyncratic part that cannot be explained by the latent factor \mathbf{f}_i . Informally, the samples $\{\mathbf{x}_i\}$ are, in some sense, assumed to be approximately embedded in a low-dimensional subspace encoded by the loading matrix \mathbf{L}^* , which describes how each component of data \mathbf{x}_i depends on the factor \mathbf{f}_i and captures the inter-dependency across different variables. In the language of PCA, the subspace spanned by \mathbf{L}^* specifies the r principal components underlying this sequence of data samples. A common goal thus amounts to estimating the subspace spanned by the loading matrix \mathbf{L}^* and the latent factors $\{\mathbf{f}_i\}$. In the PCA literature, the subspace represented by \mathbf{L}^* is commonly referred to as the principal subspace.

In this monograph, we concentrate on the following tractable statistical model for pedagogical reasons. See Fan *et al.* (2020b, Chapter 10) for more general settings (including, say, heavy-tailed distributions and non-isotropic noise covariance matrices).

Assumption 3.1. The vectors \mathbf{f}_i and $\boldsymbol{\eta}_i$ ($1 \leq i \leq n$) are all independently generated according to

$$\mathbf{f}_i \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mathbf{0}, \mathbf{I}_r), \quad \text{and} \quad \boldsymbol{\eta}_i \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_p). \quad (3.20)$$

Moreover, we assume without loss of generality that $\mathbf{L}^* = \mathbf{U}^* (\boldsymbol{\Lambda}^*)^{1/2}$, where the columns of $\mathbf{U}^* \in \mathbb{R}^{p \times r}$ are composed of orthonormal vectors, and $\boldsymbol{\Lambda}^* = \text{diag}([\lambda_1^*, \dots, \lambda_r^*])$ is an r -dimensional diagonal matrix obeying $\lambda_1^* \geq \dots \geq \lambda_r^* > 0$. Throughout this section, we denote by

$$\kappa := \lambda_1^* / \lambda_r^*$$

the condition number of the low-rank matrix $\mathbf{L}^* \mathbf{L}^{*\top} = \mathbf{U}^* \boldsymbol{\Lambda}^* \mathbf{U}^{*\top}$.

3.3.2 Algorithm

As a starting point, it is readily seen under Assumption 3.1 that

$$\mathbf{x}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{M}^*) \quad \text{with } \mathbf{M}^* := \mathbf{U}^* \boldsymbol{\Lambda}^* \mathbf{U}^{*\top} + \sigma^2 \mathbf{I}_p. \quad (3.21)$$

In brief, the covariance matrix \mathbf{M}^* is a low-rank matrix superimposed by a scaled identity matrix; for this reason, this model is also frequently referred to as the spiked covariance model (Johnstone, 2001). The key takeaway is that the top- r eigenspace of the covariance matrix \mathbf{M}^* in (3.21) coincides with the r -dimensional principal subspace being sought after (i.e., the one spanned by \mathbf{L}^* or \mathbf{U}^*).

The above observation motivates a simple spectral algorithm, which begins by computing a sample covariance matrix

$$\mathbf{M} := \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^\top, \quad (3.22)$$

followed by computation of the rank- r eigendecomposition $\mathbf{U} \mathbf{\Lambda} \mathbf{U}^\top$ of \mathbf{M} . Here, $\mathbf{\Lambda} \in \mathbb{R}^{r \times r}$ is a diagonal matrix whose diagonal entries entail the r largest eigenvalues $\lambda_1 \geq \dots \geq \lambda_r$ of \mathbf{M} , and $\mathbf{U} := [\mathbf{u}_1, \dots, \mathbf{u}_r] \in \mathbb{R}^{p \times r}$ with \mathbf{u}_i representing the eigenvector of \mathbf{M} associated with λ_i . The spectral algorithm studied herein then returns \mathbf{U} as the estimate for the principal subspace \mathbf{U}^* .

Remark 3.3. In the presence of missing data or heteroskedastic noise (meaning that the variance of the noise entries varies across different entries), the second part of the covariance matrix \mathbf{M}^* (i.e., $\sigma^2 \mathbf{I}_p$ in (3.21)) might no longer be a scaled identity. Under such circumstances, one might need to carefully adjust the diagonal entries of \mathbf{M} in order for the algorithm to succeed; see, e.g., Lounici (2014), Loh and Wainwright (2012), Zhang *et al.* (2021), Cai *et al.* (2021a), Zhu *et al.* (2019), and Yan *et al.* (2021). The reader might consult Section 3.9 for an introduction to a commonly adopted diagonal deletion idea to address the aforementioned issue.

3.3.3 Performance guarantees

This subsection develops statistical guarantees for the spectral method described above by invoking the eigenspace perturbation theory introduced previously. The first step is to establish a connection between the sample covariance \mathbf{M} and the true covariance \mathbf{M}^* . Defining $\mathbf{F} := [\mathbf{f}_1, \dots, \mathbf{f}_n] \in \mathbb{R}^{r \times n}$ and $\mathbf{Z} := [\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_n] \in \mathbb{R}^{p \times n}$, one can easily

compute that

$$\mathbf{M} = \frac{1}{n}(\mathbf{L}^* \mathbf{F} + \mathbf{Z})(\mathbf{L}^* \mathbf{F} + \mathbf{Z})^\top = \mathbf{M}^* + \mathbf{E}, \quad (3.23)$$

where \mathbf{M}^* is defined in (3.21), and

$$\begin{aligned} \mathbf{E} := \mathbf{L}^* & \left(\frac{1}{n} \mathbf{F} \mathbf{F}^\top - \mathbf{I}_r \right) \mathbf{L}^{*\top} + \frac{1}{n} \mathbf{L}^* \mathbf{F} \mathbf{Z}^\top + \frac{1}{n} \mathbf{Z} \mathbf{F}^\top \mathbf{L}^{*\top} \\ & + \left(\frac{1}{n} \mathbf{Z} \mathbf{Z}^\top - \sigma^2 \mathbf{I}_p \right). \end{aligned} \quad (3.24)$$

To apply the Davis-Kahan theorem, we are in need of controlling the size of the perturbation matrix \mathbf{E} . This is achieved by the following lemma, whose proof is deferred to Section 3.3.4.

Lemma 3.5. Consider the settings in Section 3.3.1. Suppose that $n \geq cr \log^3(n + p)$ for some sufficiently large constant $c > 0$. Then with probability exceeding $1 - O((n + p)^{-10})$, one has

$$\|\mathbf{E}\| \lesssim \left(\lambda_1^* \sqrt{\frac{r}{n}} + \sigma \sqrt{\frac{\lambda_1^* p}{n}} + \sigma^2 \sqrt{\frac{p}{n}} + \frac{\sigma^2 p \log^{\frac{3}{2}}(n + p)}{n} \right) \log^{\frac{1}{2}}(n + p).$$

With Lemma 3.5 in place, we are ready to present the following theorem that controls the estimation error of the spectral algorithm.

Theorem 3.6. Consider the settings in Section 3.3.1. Suppose that $n \geq C(\kappa^2 r + r \log^2(n + p) + \frac{\kappa \sigma^2 p}{\lambda_r^*} + \frac{\sigma^4 p}{(\lambda_r^*)^2}) \log^3(n + p)$ for some sufficiently large constant $C > 0$. Then with probability at least $1 - O((n + p)^{-10})$, the following holds:

$$\text{dist}(\mathbf{U}, \mathbf{U}^*) \lesssim \left(\frac{\sigma}{\sqrt{\lambda_r^*}} \sqrt{\frac{\kappa p}{n}} + \frac{\sigma^2}{\lambda_r^*} \sqrt{\frac{p}{n}} + \kappa \sqrt{\frac{r}{n}} \right) \log^{\frac{1}{2}}(n + p). \quad (3.25)$$

Remark 3.4. The third term $\kappa \sqrt{(r \log(n + p))/n}$ on the right-hand side of (3.25) arises due to the randomness of $\{\mathbf{f}_i\}$ but not that of $\{\boldsymbol{\eta}_i\}$. If our goal is instead to estimate the eigenspace of $\mathbf{L}^* (\frac{1}{n} \sum_i \mathbf{f}_i \mathbf{f}_i^\top) \mathbf{L}^{*\top}$ as opposed to that of $\mathbf{L}^* \mathbf{L}^{*\top}$, then this term can be erased.

To interpret what Theorem 3.6 conveys, we include a few remarks in the sequel, focusing on the simple scenario where $\kappa = O(1)$. In view of Remark 3.4, we shall ignore the term $\kappa \sqrt{(r \log(n + p))/n}$ in the discussion below.

Linear vs. quadratic dependency on the noise level. In comparison to the matrix denoising task (cf. Section 3.2.2) where $\text{dist}(\mathbf{U}, \mathbf{U}^*)$ scales linearly with the noise level σ (cf. (3.13)), the above performance guarantees for PCA exhibit contrasting behavior in two different regimes depending on the strength of the signal-to-noise ratio (SNR), measured in terms of λ_r^*/σ^2 :

- When the SNR is sufficiently large with $\lambda_r^*/\sigma^2 \gtrsim 1$, then the dominant factor in (3.25) is the term $\sigma \left(\frac{p \log(n+p)}{\lambda_r^* n} \right)^{1/2}$, which scales linearly with the noise level.
- When the SNR drops below the threshold $\lambda_r^*/\sigma^2 \lesssim 1$, then the term $\frac{\sigma^2}{\lambda_r^*} \sqrt{\frac{p \log(n+p)}{n}}$ —which scales quadratically with the noise level—enters the picture and becomes the dominant effect.

In truth, the quadratic term emerges since our spectral method operates upon the sample covariance matrix, which inevitably contains second moments of the noise components.

Tightness and optimality. Natural questions arise as to whether the performance guarantees in Theorem 3.6 are tight, and whether the statistical accuracy can be further improved by designing more intelligent algorithms. These questions can be addressed by looking into the fundamental statistical limits. As established in the literature (Zhang *et al.*, 2021; Cai *et al.*, 2021a), one cannot hope to achieve

$$\text{dist}(\widehat{\mathbf{U}}, \mathbf{U}^*) = o\left(\frac{\sigma}{\sqrt{\lambda_r^*}} \sqrt{\frac{p}{n}} + \frac{\sigma^2}{\lambda_r^*} \sqrt{\frac{p}{n}}\right) \quad (3.26)$$

in a minimax sense, regardless of the choice of the estimator $\widehat{\mathbf{U}}$; see, e.g., Zhang *et al.* (2021, Theorem 2) for a precise statement. Comparing (3.26) with Theorem 3.6 reveals the near statistical optimality of the spectral method (modulo some log factor), and confirms the tightness of the eigenspace perturbation theory when applied to this problem.

Proof of Theorem 3.6. We first make the observation that

$$\lambda_1(\mathbf{M}^*) \geq \dots \geq \lambda_r(\mathbf{M}^*) > \lambda_{r+1}(\mathbf{M}^*) = \dots = \lambda_p(\mathbf{M}^*) = \sigma^2 > 0,$$

$$\text{and } \lambda_r(\mathbf{M}^*) - \lambda_{r+1}(\mathbf{M}^*) = \lambda_r^*.$$

The Davis-Kahan $\sin\Theta$ theorem (cf. Corollary 2.8) thus implies that: if the perturbation size obeys $\|\mathbf{E}\| \leq (1 - 1/\sqrt{2})\lambda_r^*$, then one has

$$\begin{aligned} \text{dist}(\mathbf{U}, \mathbf{U}^*) &\leq \frac{2\|\mathbf{E}\|}{\lambda_r(\mathbf{M}^*) - \lambda_{r+1}(\mathbf{M}^*)} = \frac{2\|\mathbf{E}\|}{\lambda_r^*} \\ &\lesssim \frac{1}{\lambda_r^*} \left(\lambda_1^* \sqrt{\frac{r}{n}} + \sigma \sqrt{\frac{\lambda_1^* p}{n}} + \sigma^2 \sqrt{\frac{p}{n}} + \frac{\sigma^2 p \log^{\frac{3}{2}}(n+p)}{n} \right) \log^{\frac{1}{2}}(n+p) \\ &\asymp \left(\kappa \sqrt{\frac{r}{n}} + \frac{\sigma}{\sqrt{\lambda_r^*}} \sqrt{\frac{\kappa p}{n}} + \frac{\sigma^2}{\lambda_r^*} \sqrt{\frac{p}{n}} \right) \log^{\frac{1}{2}}(n+p). \end{aligned}$$

Here, the penultimate inequality results from Lemma 3.5; the last line is valid as long as $n \gtrsim (\sigma^2/\lambda_1^*)p \log^3(n+p)$ —a condition that would hold under the assumption of this theorem—so that the fourth term is dominated by the second one in the parenthesis of the penultimate line. Finally, it is immediately seen from Lemma 3.5 that the condition $\|\mathbf{E}\| \leq (1 - 1/\sqrt{2})\lambda_r^*$ would hold under the assumption of this theorem.

3.3.4 Proof of Lemma 3.5

We start by applying the triangle inequality to (3.24) as follows

$$\begin{aligned} \|\mathbf{E}\| &\leq \|\mathbf{L}^*\|^2 \left\| \frac{1}{n} \mathbf{F} \mathbf{F}^\top - \mathbf{I}_r \right\| + \|\mathbf{L}^*\| \left\| \frac{1}{n} \mathbf{F} \mathbf{Z}^\top \right\| + \left\| \frac{1}{n} \mathbf{Z} \mathbf{F}^\top \right\| \|\mathbf{L}^*\| \\ &\quad + \left\| \frac{1}{n} \mathbf{Z} \mathbf{Z}^\top - \sigma^2 \mathbf{I}_p \right\|. \end{aligned} \tag{3.27}$$

In order to develop an upper bound on this quantity, one needs to control the spectral norm of $\frac{1}{n} \mathbf{F} \mathbf{F}^\top - \mathbf{I}_r$, $\frac{1}{n} \mathbf{F} \mathbf{Z}^\top$, $\frac{1}{n} \mathbf{Z} \mathbf{F}^\top$ and $\frac{1}{n} \mathbf{Z} \mathbf{Z}^\top - \sigma^2 \mathbf{I}_p$. All of these terms share similar randomness structure, namely, they are all averages of independent zero-mean random matrices. As a result, the truncated matrix Bernstein inequality in Corollary 3.2 becomes applicable. In what follows, we shall only demonstrate how to control the size of $\frac{1}{n} \mathbf{F} \mathbf{Z}^\top$; the other terms can be bounded similarly.

Write $\mathbf{F} \mathbf{Z}^\top = \sum_{i=1}^n \mathbf{f}_i \boldsymbol{\eta}_i^\top$. Since the entries of $\mathbf{F} \mathbf{Z}^\top$ might be unbounded, we start by identifying an appropriate truncation level. From standard properties about Gaussian distributions and the union bound, it is straightforward to verify that

$$\mathbb{P} \left\{ \|\mathbf{f}_i\|_\infty \leq 5\sqrt{\log(n+p)} \text{ and } \|\boldsymbol{\eta}_i\|_\infty \leq 5\sigma\sqrt{\log(n+p)} \right\} \geq 1 - (n+p)^{-11.5}.$$

One can further derive

$$\|\mathbf{f}_i \boldsymbol{\eta}_i^\top\| \leq \|\mathbf{f}_i\|_2 \|\boldsymbol{\eta}_i\|_2 \leq \sqrt{rp} \|\mathbf{f}_i\|_\infty \|\boldsymbol{\eta}_i\|_\infty \leq 25\sqrt{rp}\sigma \log(n+p)$$

with probability greater than $1 - (n+p)^{-11.5}$. In other words, with the choice $L := 25\sqrt{rp}\sigma \log(n+p)$ one has

$$\mathbb{P}\left\{\|\mathbf{f}_i \boldsymbol{\eta}_i^\top\| \geq L\right\} \leq (n+p)^{-11.5} =: q_0.$$

Additionally, the symmetry of Gaussian distributions implies

$$\mathbb{E}[\mathbf{f}_i \boldsymbol{\eta}_i^\top] - \mathbb{E}\left[\mathbf{f}_i \boldsymbol{\eta}_i^\top \mathbb{1}\{\|\mathbf{f}_i \boldsymbol{\eta}_i^\top\| < L\}\right] = 0.$$

To invoke the truncated Bernstein inequality, it remains to determine the variance statistic. Towards this end, letting $\mathbf{B}_i = \mathbf{f}_i \boldsymbol{\eta}_i^\top$, we observe that

$$\begin{aligned}\mathbb{E}[\mathbf{B}_i \mathbf{B}_i^\top] &= \mathbb{E}[\mathbf{f}_i \boldsymbol{\eta}_i^\top \boldsymbol{\eta}_i \mathbf{f}_i^\top] = \mathbb{E}[\boldsymbol{\eta}_i^\top \boldsymbol{\eta}_i] \mathbb{E}[\mathbf{f}_i \mathbf{f}_i^\top] = p\sigma^2 \mathbf{I}_r, \\ \mathbb{E}[\mathbf{B}_i^\top \mathbf{B}_i] &= \mathbb{E}[\boldsymbol{\eta}_i \mathbf{f}_i^\top \mathbf{f}_i \boldsymbol{\eta}_i^\top] = \mathbb{E}[\mathbf{f}_i^\top \mathbf{f}_i] \mathbb{E}[\boldsymbol{\eta}_i \boldsymbol{\eta}_i^\top] = r\sigma^2 \mathbf{I}_p,\end{aligned}$$

thus leading to

$$v := \max\left\{\left\|\sum_i \mathbb{E}[\mathbf{B}_i \mathbf{B}_i^\top]\right\|, \left\|\sum_i \mathbb{E}[\mathbf{B}_i^\top \mathbf{B}_i]\right\|\right\} = np\sigma^2,$$

where we use the fact that $r \leq p$. Taking these bound together and applying the truncated matrix Bernstein theorem (see Corollary 3.2) demonstrate that if $n \gtrsim r \log^3(n+p)$, one has

$$\begin{aligned}\frac{1}{n} \|\mathbf{F} \mathbf{Z}^\top\| &\lesssim \frac{1}{n} \sqrt{v \log(n+p)} + \frac{1}{n} L \log(n+p) \\ &\asymp \sigma \sqrt{\frac{p \log(n+p)}{n}} + \frac{\sqrt{rp}}{n} \sigma \log^2(n+p) \asymp \sigma \sqrt{\frac{p \log(n+p)}{n}}\end{aligned}\quad (3.28a)$$

with probability at least $1 - O((n+p)^{-10}) - nq_0 = 1 - O((n+p)^{-10})$.

Repeating the above analysis yields that: if $n \gtrsim r \log^3(n+p)$, with probability at least $1 - O((n+p)^{-10})$ one has

$$\left\|\frac{1}{n} \mathbf{F} \mathbf{F}^\top - \mathbf{I}_r\right\| \lesssim \sqrt{\frac{r \log(n+p)}{n}},\quad (3.28b)$$

$$\left\|\frac{1}{n} \mathbf{Z} \mathbf{Z}^\top - \sigma^2 \mathbf{I}_p\right\| \lesssim \sigma^2 \sqrt{\frac{p \log(n+p)}{n}} + \frac{\sigma^2 p \log^2(n+p)}{n}.\quad (3.28c)$$

Note that we do not get rid of the second term on the right-hand side of (3.28c) since we do not assume $n \gtrsim p \log^3(n + p)$.

Substituting the above results (3.28) into (3.27) and recognizing the basic fact $\|\mathbf{L}^*\| = \|\mathbf{U}^*(\mathbf{\Lambda}^*)^{1/2}\| \leq \|(\mathbf{\Lambda}^*)^{1/2}\| = \sqrt{\lambda_1^*}$, we conclude that

$$\|\mathbf{E}\| \lesssim \left(\lambda_1^* \sqrt{\frac{r}{n}} + \sigma \sqrt{\frac{\lambda_1^* p}{n}} + \sigma^2 \sqrt{\frac{p}{n}} + \frac{\sigma^2 p \log^{\frac{3}{2}}(n + p)}{n} \right) \log^{\frac{1}{2}}(n + p).$$

3.4 Graph clustering and community recovery

Next, we move on to a central problem that permeates data science applications: clustering. An important formulation that falls under this category is graph clustering or community recovery, which aims to cluster individuals into different communities based on pairwise measurements of their relationships, each of which reveals information about whether or not two individuals belong to the same community (Abbe, 2017); see Figure 3.1 for an illustration. There has been a recent explosion of interest in this problem, due to its wide applicability in, say, social network analysis (Azaouzi *et al.*, 2019), image segmentation (Browet *et al.*, 2011), shape mapping in computer vision (Huang and Guibas, 2013), haplotype phasing in genome sequencing (Chen *et al.*, 2016a), to name just a few. This section explores the capability of spectral methods in application to graph clustering; we will revisit the clustering problem again in Section 3.5 for another common formulation.

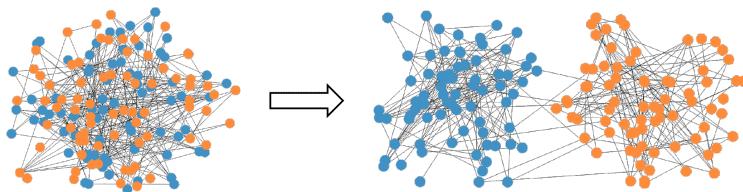


Figure 3.1: Illustration of graph clustering and community recovery, where one wishes to cluster all nodes into two communities based on the edges in the graph.

3.4.1 Problem formulation and assumptions

In this section, we formulate the graph clustering problem via the well-renowned *stochastic block model* (SBM) introduced in Holland *et al.* (1983)—an idealized generative model that commonly serves as a theoretical benchmark for evaluating community recovery algorithms.

Consider an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ that comprises n vertices, where \mathcal{V} and \mathcal{E} denote the vertex set and the edge set of \mathcal{G} , respectively. The n vertices, labelled by $1, \dots, n$, exhibit community structures and can be grouped into two non-overlapping communities of *equal* sizes. Here and throughout, n is assumed to be an even number, so that each community contains exactly $n/2$ vertices. To encode the community memberships, we assign n binary-valued variables $x_i^* \in \{1, -1\}$ ($1 \leq i \leq n$) to the vertices in a way that

$$x_i^* = \begin{cases} 1, & \text{if vertex } i \text{ belongs to the 1st community,} \\ -1, & \text{otherwise.} \end{cases}$$

The SBM assumes that the set \mathcal{E} of (undirected) edges is generated randomly based on the community memberships of the incident vertices. To be precise, each pair (i, j) of vertices is connected by an edge independently with probability p (resp. q) if i and j belong to the same community (resp. different communities). The resultant connectivity pattern is represented by an adjacency matrix $\mathbf{A} = [A_{i,j}]_{1 \leq i,j \leq n} \in \{0, 1\}^{n \times n}$, such that for each pair (i, j) ,

$$A_{i,j} = \begin{cases} 1, & \text{if } (i, j) \in \mathcal{E}, \\ 0, & \text{otherwise.} \end{cases} \quad (3.29)$$

By convention, we take the diagonal entries to be $A_{i,i} = 0$ for all $1 \leq i \leq n$. As a remark, the matrix \mathbf{A} is symmetric since \mathcal{G} is an undirected graph, with upper triangular elements being realizations of independent Bernoulli random variables with mean either p (if two nodes are in the same community) or q (otherwise). In addition, it is assumed throughout that $p > q > 0$, implying that there are in expectation more within-community edges than across-community edges.

Based on the adjacency matrix \mathbf{A} generated by the SBM, the goal is to identify the latent community memberships of the vertices. To

phrase it in mathematical terms, the aim is to reconstruct the vector $\mathbf{x}^* = [x_i^*]_{1 \leq i \leq n} \in \{1, -1\}^n$ modulo the global sign, namely, recovering either \mathbf{x}^* or $-\mathbf{x}^*$. This is all one can hope for, as there is absolutely no basis to distinguish the names of two groups.

3.4.2 Algorithm: spectral clustering

Now we describe a spectral method. To simplify presentation, it is assumed without loss of generality that: $x_i^* = 1$ for any $1 \leq i \leq n/2$, and $x_i^* = -1$ for any $i > n/2$.

A starting point for the algorithm design is to examine the mean of the adjacency matrix, given as follows

$$\mathbb{E}[\mathbf{A}] = \begin{bmatrix} p \mathbf{1}_{n/2} \mathbf{1}_{n/2}^\top & q \mathbf{1}_{n/2} \mathbf{1}_{n/2}^\top \\ q \mathbf{1}_{n/2} \mathbf{1}_{n/2}^\top & p \mathbf{1}_{n/2} \mathbf{1}_{n/2}^\top \end{bmatrix} - p\mathbf{I}.$$

As revealed by the above calculation, the matrix constructed below

$$\mathbf{M} = \mathbf{A} - \frac{p+q}{2} \mathbf{1}_n \mathbf{1}_n^\top + p\mathbf{I} \quad (3.30)$$

exhibits an approximate rank-1 structure, in the sense that its mean

$$\mathbf{M}^* := \mathbb{E}[\mathbf{M}] = \frac{p-q}{2} \begin{bmatrix} \mathbf{1}_{n/2} \\ -\mathbf{1}_{n/2} \end{bmatrix} \begin{bmatrix} \mathbf{1}_{n/2}^\top & -\mathbf{1}_{n/2}^\top \end{bmatrix} \quad (3.31)$$

is a rank-1 matrix. The leading eigenvalue of \mathbf{M}^* and its associated eigenvector are given respectively by

$$\lambda^* := \frac{(p-q)n}{2}, \quad \text{and} \quad \mathbf{u}^* := \frac{1}{\sqrt{n}} \begin{bmatrix} \mathbf{1}_{n/2} \\ -\mathbf{1}_{n/2} \end{bmatrix}. \quad (3.32)$$

Crucially, the eigenvector \mathbf{u}^* encapsulates the precise community structure we seek to recover: all positive entries of \mathbf{u}^* correspond to vertices from one community, while the remaining ones form another community.

Inspired by the above calculation, a candidate spectral clustering algorithm consists of eigendecomposition followed by entrywise rounding:

1. Compute the leading eigenvector \mathbf{u} of \mathbf{M} (constructed in (3.30));
2. Compute the estimate $\mathbf{x} = [x_i]_{1 \leq i \leq n}$ such that for any $1 \leq i \leq n$,

$$x_i = \text{sgn}(u_i) = \begin{cases} 1, & \text{if } u_i > 0, \\ -1, & \text{if } u_i \leq 0. \end{cases} \quad (3.33)$$

In words, the community memberships are estimated in accordance with the signs of the entries of the leading eigenvector of \mathbf{M} , namely, the entries with the same signs are declared to come from the same cluster.

Remark 3.5. The above algorithm requires prior knowledge of the parameters p and q when constructing \mathbf{M} . It is also feasible to develop a “model-agnostic” alternative by, for instance, looking at the second eigenvector of \mathbf{A} (since the second eigenvector of $\mathbb{E}[\mathbf{A}]$ turns out to be precisely \mathbf{u}^*), which does not rely on prior information about p and q at all; see, e.g., Abbe *et al.* (2020b) for details. Here, we adopt the above model-dependent version primarily for convenience of exposition.

3.4.3 Performance guarantees: almost exact recovery

The spectral method enjoys appealing statistical guarantees for recovering the community structure of the SBM, which can be readily obtained by invoking the ℓ_2 eigenvector perturbation theory. To demonstrate this, we begin by developing an upper bound on the spectral norm of the perturbation matrix $\mathbf{E} := \mathbf{M} - \mathbf{M}^*$, postponing the proof to Section 3.4.4.

Lemma 3.7. Consider the settings in Section 3.4.1, and suppose that $np \gtrsim \log n$. Then with probability at least $1 - O(n^{-8})$, one has

$$\|\mathbf{E}\| \lesssim \sqrt{np}. \quad (3.34)$$

This spectral norm bound, in conjunction with the Davis-Kahan $\sin\Theta$ theorem, leads to the following theoretical support for the spectral method introduced in Section 3.4.2.

Theorem 3.8. Consider the setting in Section 3.4.1, and suppose that

$$p \gtrsim \frac{\log n}{n}, \quad \text{and} \quad \sqrt{\frac{p}{n}} = o(p - q). \quad (3.35)$$

With probability exceeding $1 - O(n^{-8})$, the spectral method achieves

$$\frac{1}{n} \sum_{i=1}^n \mathbb{1}\{x_i = x_i^*\} = 1 - o(1), \quad \text{or} \quad \frac{1}{n} \sum_{i=1}^n \mathbb{1}\{x_i = -x_i^*\} = 1 - o(1).$$

It is noteworthy that the metric

$$\min \left\{ \frac{1}{n} \sum_{i=1}^n \mathbb{1}\{x_i \neq x_i^*\}, \frac{1}{n} \sum_{i=1}^n \mathbb{1}\{x_i \neq -x_i^*\} \right\}$$

can be understood as the mis-clustering rate. In a nutshell, Theorem 3.8 asserts that with the assistance of simple rounding (i.e., the $\text{sgn}(\cdot)$ operation), the spectral method allows for *almost* exact community recovery—namely, correctly clustering all but a vanishing fraction of the vertices—assuming satisfaction of Condition (3.35). Note that “almost exact recovery” is also referred to as “weak consistency” in the literature (Abbe, 2017).

Let us take a moment to interpret the recovery condition in (3.35). The first requirement in Condition (3.35) ensures the presence of sufficiently many edges in the observed graph, while still permits the graph to be fairly sparse (with average vertex degrees as low as the order of $\log n$). The second requirement in Condition (3.35)—which imposes a lower bound on the separation between the edge densities p and q —guarantees that the within-community edges can be adequately differentiated from across-community edges. As a more concrete example, consider the scenario where $p \asymp (\log n)/n$ (so that each vertex is only expected to be incident to $O(\log n)$ edges). In this case, the second requirement in Condition (3.35) can be translated into

$$p - q \gg \sqrt{\log n} / n, \quad \text{if } p \asymp (\log n) / n.$$

This indicates that the separation $p - q$ is allowed to be considerably smaller than the edge densities, even in this low-edge-density regime. In comparison, in another extreme case with $p \asymp 1$ (so that each vertex is likely to be connected with a constant fraction of other vertices), the second requirement in Condition (3.35) reads

$$p - q \gg 1/\sqrt{n}, \quad \text{if } p \asymp 1,$$

thereby allowing the edge density difference to be even \sqrt{n} times smaller than the edge densities themselves.

It is worth highlighting that the spectral method is not merely capable of correctly clustering all but a diminishing fraction of vertices;

in fact, it allows for simultaneous and exact recovery for *all* vertices under slightly modified conditions. Establishing this stronger assertion requires developing a significantly strengthened ℓ_∞ -based eigenvector perturbation theory, which will be elucidated in Section 4.5. The discussion about the statistical optimality of this spectral method is postponed to Section 4.5 as well.

Proof of Theorem 3.8. It is readily seen from Lemma 3.7 that with probability at least $1 - O(n^{-8})$,

$$\|\mathbf{E}\| \leq \left(1 - \frac{1}{\sqrt{2}}\right) \frac{n(p-q)}{2} = \left(1 - \frac{1}{\sqrt{2}}\right) \lambda^*,$$

provided that Condition (3.35) holds. Here, λ^* is defined in (3.32). Apply Corollary 2.8 to yield that with probability at least $1 - O(n^{-8})$,

$$\text{dist}(\mathbf{u}, \mathbf{u}^*) \leq \frac{2\|\mathbf{E}\|}{\lambda^*} \lesssim \frac{\sqrt{np}}{n(p-q)} = o(1), \quad (3.36)$$

where the last relation follows from Condition (3.35).

Assume, without loss of generality, that $\|\mathbf{u} - \mathbf{u}^*\|_2 = \text{dist}(\mathbf{u}, \mathbf{u}^*)$. We shall pay attention to the set

$$\mathcal{N} := \{i \mid |u_i - u_i^*| \geq 1/\sqrt{n}\}.$$

In view of the rounding procedure: for any i obeying $x_i \neq x_i^*$, one necessarily has $\text{sgn}(u_i) \neq \text{sgn}(u_i^*)$, thus indicating that $|u_i - u_i^*| \geq |u_i^*| = 1/\sqrt{n}$ and hence $i \in \mathcal{N}$. Combining the ℓ_2 bound (3.36) and the definition of \mathcal{N} , we can easily verify that

$$|\mathcal{N}| \leq \frac{\|\mathbf{u} - \mathbf{u}^*\|_2^2}{(1/\sqrt{n})^2} = o(n),$$

which in turn leads to the advertised result

$$\frac{1}{n} \sum_{i=1}^n \mathbb{1}\{x_i \neq x_i^*\} \leq \frac{1}{n} \sum_{i=1}^n \mathbb{1}\left\{ |u_i - u_i^*| \geq \frac{1}{\sqrt{n}} \right\} = \frac{|\mathcal{N}|}{n} = o(1).$$

3.4.4 Proof of Lemma 3.7

We intend to apply Theorem 3.4 to establish this lemma. First, observe from the definition $\mathbf{E} = \mathbf{M} - \mathbf{M}^* = \mathbf{A} - \mathbb{E}[\mathbf{A}]$ that

$$|E_{i,j}| \leq \max_{i,j} |A_{i,j}| = 1.$$

In addition, the variance of $E_{i,j}$ is upper bounded by

$$\mathbb{E}[E_{i,j}^2] = \text{Var}(A_{i,j}) \leq \mathbb{E}[A_{i,j}^2] \stackrel{\text{(i)}}{\leq} \max\{p, q\} \stackrel{\text{(ii)}}{=} p$$

for any (i, j) , where (i) follows since $A_{i,j}$ is a Bernoulli random variable with mean either p or q , and (ii) is due to the assumption $p > q$. The bound (3.9) and the condition $np \gtrsim \log n$ thus imply that

$$\|\mathbf{E}\| = \|\mathbf{M} - \mathbb{E}[\mathbf{M}]\| \lesssim \sqrt{np} + O(\sqrt{\log n}) \asymp \sqrt{np} \quad (3.37)$$

with probability exceeding $1 - O(n^{-8})$.

3.5 Clustering in Gaussian mixture models

This section is also concerned with clustering, with the aim of grouping unlabeled data points into a few clusters (so that the data within the same cluster share similar characteristics). In contrast to the graph clustering setting in Section 3.4 where only pairwise measurements are available, this section assumes direct access to data samples for each individual. Spectral methods—possibly with the aid of subsequent refinement like k -means—continue to be remarkably effective for this setting, achieving practical success in, say, image segmentation (Shi and Malik, 2000), text separation (Reynolds and Rose, 1995), climate modeling (Lin *et al.*, 2017), and heterogeneity modeling in precision medicine and marketing (Fan *et al.*, 2014). Motivated by the empirical successes, understanding the theoretical properties of spectral clustering has garnered growing attention recently. In particular, Gaussian mixture models emerge as a succinct model of attack, providing elegant yet intuitive abstractions to pivotal quantities that dictate the feasibility of spectral clustering.

3.5.1 Gaussian mixture models and assumptions

Model and goal. Imagine that we have collected n independent samples $\{\mathbf{x}_i\}_{1 \leq i \leq n}$, generated from a mixture of r spherical Gaussians with respective centers $\boldsymbol{\theta}_1^*, \dots, \boldsymbol{\theta}_r^* \in \mathbb{R}^p$. More precisely, for each sample vector $\mathbf{x}_i \in \mathbb{R}^p$, we assume the existence of a predetermined, yet *a priori*

unknown, cluster membership variable $\xi_i^* \in [r]$ such that

$$\mathbf{x}_i = \begin{cases} \boldsymbol{\theta}_1^* + \boldsymbol{\eta}_i, & \text{if } \xi_i^* = 1, \\ \vdots & \vdots \\ \boldsymbol{\theta}_r^* + \boldsymbol{\eta}_i, & \text{if } \xi_i^* = r, \end{cases} \quad (3.38)$$

where the noise vector $\boldsymbol{\eta}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_p)$ is independently generated across the samples. In words, ξ_i^* indicates which Gaussian component a sample is generated from. Clustering in this Gaussian mixture model can, therefore, be posed as recovering the set of cluster membership variables $\{\xi_i^*\}_{1 \leq i \leq n}$ (modulo the global permutation ambiguity).

Assumptions. To simplify our exposition, we impose the following assumptions throughout this section. As a worthy note, this assumption is often non-essential and can be significantly relaxed, which we shall remark on momentarily in Remark 3.8.

Assumption 3.2. The centers are independently generated obeying

$$\boldsymbol{\theta}_i^* \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}\left(\mathbf{0}, \frac{\Delta^2}{2p} \mathbf{I}_p\right), \quad 1 \leq i \leq r$$

for some parameter $\Delta > 0$.

Under this assumption, standard Gaussian concentration inequalities (Vershynin, 2017) tell us that, with high probability (for large p and $n = \text{poly}(p)$),

$$\begin{aligned} \|\boldsymbol{\theta}_i^*\|_2^2 &= (1 + o(1)) \frac{\Delta^2}{2}, & |\boldsymbol{\theta}_i^*{}^\top \boldsymbol{\theta}_j^*| &= o(\Delta^2), \\ \|\boldsymbol{\theta}_i^* - \boldsymbol{\theta}_j^*\|_2^2 &= \|\boldsymbol{\theta}_i^*\|_2^2 + \|\boldsymbol{\theta}_j^*\|_2^2 - 2\boldsymbol{\theta}_i^*{}^\top \boldsymbol{\theta}_j^* = (1 + o(1))\Delta^2 \end{aligned}$$

hold for any pair $i \neq j$, where $o(1)$ denotes a vanishingly small quantity as p approaches infinity. The indication is that the parameter Δ reflects (approximately) the separation between any pair of centers.

For simplicity of presentation, it is further assumed that there are exactly n/r samples drawn from each of the r Gaussian components.

Without loss of generality, we assume that

$$\xi_i^* = l, \quad \text{if } \left\lceil \frac{i}{r} \right\rceil = l \quad (3.39)$$

for any $1 \leq i \leq n$, where the ceiling function $\lceil x \rceil$ represents the least integer greater than or equal to the number $x \in \mathbb{R}$. In other words, the first batch of n/r samples is drawn from the first Gaussian component, the second batch comes from the second component, and so on. It is worth pointing out that this assumed assignment information (3.39) is unavailable when running the spectral clustering algorithm.

3.5.2 Algorithm and rationale

Motivation: spectral structure of the data matrix. In order to develop a spectral clustering algorithm, it is instrumental to first examine the spectral feature of the following data matrix

$$\mathbf{X} := [\mathbf{x}_1, \dots, \mathbf{x}_n] = \mathbb{E}[\mathbf{X}] + \underbrace{[\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_n]}_{=: \mathbf{Z}}. \quad (3.40)$$

Clearly, $\mathbb{E}[\mathbf{X}] \in \mathbb{R}^{p \times n}$ exhibits a rank- r structure:

$$\mathbb{E}[\mathbf{X}] = [\boldsymbol{\theta}_1^*, \dots, \boldsymbol{\theta}_1^*, \boldsymbol{\theta}_2^*, \dots, \boldsymbol{\theta}_2^*, \dots, \boldsymbol{\theta}_r^*, \dots, \boldsymbol{\theta}_r^*] = \boldsymbol{\Theta}^* \mathbf{F}^{*\top},$$

where we define

$$\boldsymbol{\Theta}^* := [\boldsymbol{\theta}_1^*, \dots, \boldsymbol{\theta}_r^*] \in \mathbb{R}^{p \times r}, \quad \mathbf{F}^* := \begin{bmatrix} \mathbf{1}_{\frac{n}{r}} & & & \\ & \mathbf{1}_{\frac{n}{r}} & & \\ & & \ddots & \\ & & & \mathbf{1}_{\frac{n}{r}} \end{bmatrix} \in \mathbb{R}^{n \times r}. \quad (3.41)$$

Similarly, the Gram matrix $\mathbf{X}^\top \mathbf{X}$ also inherits this rank- r structure in the following sense (albeit in the form of a “spiked” structure due to the presence of noise):

$$\mathbb{E}[\mathbf{X}^\top \mathbf{X}] = \mathbb{E}[\mathbf{X}]^\top \mathbb{E}[\mathbf{X}] + \mathbb{E}[\mathbf{Z}^\top \mathbf{Z}] = \mathbf{F}^* \boldsymbol{\Theta}^{*\top} \boldsymbol{\Theta}^* \mathbf{F}^{*\top} + p\mathbf{I}_n. \quad (3.42)$$

Recognizing that \mathbf{F}^* encodes all the cluster membership information, one is motivated to attempt information extraction from the rank- r eigenspace of $\mathbf{X}^\top \mathbf{X}$, akin to the PCA algorithm introduced in Section 3.3.2.

Algorithm: spectral clustering followed by k -means. With the preceding spectral properties in mind, we are ready to present a spectral clustering algorithm tailored to this Gaussian mixture model. Given that the eigenspace of $\mathbf{X}^\top \mathbf{X}$ might only approximate \mathbf{F}^* up to global rotation, we include a follow-up k -means scheme (MacQueen, 1967) to produce a valid clustering outcome based on the spectral estimate.

1. Compute the leading rank- r eigenspace $\mathbf{U} \in \mathbb{R}^{n \times r}$ of $\mathbf{X}^\top \mathbf{X}$.
2. Compute $\mathbf{Y} = \mathcal{P}(\mathbf{U}\mathbf{U}^\top) \in \mathbb{R}^{n \times n}$, where the operator $\mathcal{P}(\cdot)$ projects each column onto the unit sphere, i.e.,

$$\mathcal{P}(\mathbf{Z}) := \left[\frac{\mathbf{z}_1}{\|\mathbf{z}_1\|_2}, \dots, \frac{\mathbf{z}_n}{\|\mathbf{z}_n\|_2} \right]$$

for any matrix $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_n]$. As will be discussed below, the projection step is not necessary, and we can also simply take $\mathbf{Y} = \mathbf{U}\mathbf{U}^\top$.

3. Let \mathbf{y}_i represent the i -th column of \mathbf{Y} , and apply the k -means algorithm (with $k = r$) to the vectors $\{\mathbf{y}_i\}_{1 \leq i \leq n}$ to find the cluster centers and cluster labels for all individuals; namely, we compute

$$\left(\{\hat{\xi}_i\}_{i=1}^n, \{\hat{\vartheta}_i\}_{i=1}^r \right) = \arg \min_{\xi_1, \dots, \xi_n \in [r], \vartheta_1, \dots, \vartheta_r \in \mathbb{R}^n} \sum_{i=1}^n \|\mathbf{y}_i - \vartheta_{\xi_i}\|_2^2. \quad (3.43)$$

The algorithm then returns $\{\hat{\xi}_i\}_{1 \leq i \leq n}$ as the clustering result. Interestingly, Step 1 bears similarity with the spectral algorithm for graph clustering, since we essentially generate a pairwise similarity measurement for each pair (i, j) using the inner product $\langle \mathbf{x}_i, \mathbf{x}_j \rangle$.

Remark 3.6. The k -means formulation (3.43)—which minimizes the sum of squared distance between each data point and the center of its associated cluster—is an integer program and intractable in general (Aloise *et al.*, 2009). Fortunately, computationally feasible solutions are available either under sufficient minimum center separation or when suitably initialized (Lloyd, 1982; Vempala and Wang, 2004; Lu and Zhou, 2016; Peng and Wei, 2007; Awasthi *et al.*, 2015; Mixon *et al.*, 2017; Iguchi *et al.*, 2017). An in-depth account of this computational aspect is beyond the scope of this monograph, and the interested reader is referred to Li *et al.* (2020b) and Löffler *et al.* (2021) for details.

Further explanations. We take a moment to explain why k -means is applied to the columns of \mathbf{Y} . Recall that the central object the spectral algorithm seeks to approximate is the leading rank- r eigenspace of $\mathbf{F}^* \Theta^{*\top} \Theta^* \mathbf{F}^{*\top}$ (cf. (3.42)). For convenience, suppose we have the eigen-decomposition $\Theta^{*\top} \Theta^* = \mathbf{U}_\theta \Sigma_\theta \mathbf{U}_\theta^\top$, where $\mathbf{U}_\theta \in \mathcal{O}^{r \times r}$ is orthonormal and $\Sigma_\theta \in \mathbb{R}^{r \times r}$ is diagonal. This results in the decomposition

$$\mathbf{F}^* \Theta^{*\top} \Theta^* \mathbf{F}^{*\top} = \frac{n}{r} \cdot \underbrace{\left(\sqrt{\frac{r}{n}} \mathbf{F}^* \mathbf{U}_\theta \right)}_{=: \mathbf{U}^*} \Sigma_\theta \left(\sqrt{\frac{r}{n}} \mathbf{F}^* \mathbf{U}_\theta \right)^\top. \quad (3.44)$$

Apparently, the matrix $\mathbf{U}^* \in \mathbb{R}^{n \times r}$ defined above has orthonormal columns and, as a result, represents the eigenspace of $\mathbf{F}^* \Theta^{*\top} \Theta^* \mathbf{F}^{*\top}$. The idea is that if the spectral estimate \mathbf{U} approximates $\mathbf{U}^* \in \mathbb{R}^{n \times r}$ well, then the matrices $\sqrt{\frac{n}{r}} \mathbf{U} \mathbf{U}^\top$ and \mathbf{Y} constructed above are hopefully close to the following matrix

$$\mathbf{Y}^* := \sqrt{\frac{n}{r}} \mathbf{U}^* \mathbf{U}^{*\top} = \sqrt{\frac{r}{n}} \begin{bmatrix} \mathbf{1}_{\frac{n}{r}} \mathbf{1}_{\frac{n}{r}}^\top & & \\ & \ddots & \\ & & \mathbf{1}_{\frac{n}{r}} \mathbf{1}_{\frac{n}{r}}^\top \end{bmatrix}. \quad (3.45)$$

As can be easily seen, the data points belonging to the same ground-truth cluster are associated with identical columns in \mathbf{Y}^* ; for instance, each of the first n/r samples—which belongs to the first cluster—corresponds to a column of \mathbf{Y}^* given by $\sqrt{\frac{r}{n}} \begin{bmatrix} \mathbf{1}_{n/r} \\ \mathbf{0} \end{bmatrix}$. Therefore, clustering the columns of \mathbf{Y} via k -means is expected to unveil the underlying cluster structure, provided that \mathbf{Y} is sufficiently close to \mathbf{Y}^* . In summary, spectral estimation (Steps 1-2) effectively leads to a new vector \mathbf{y}_i for each point, which enjoys substantially enhanced signal-to-noise ratio compared to \mathbf{x}_i and boosts the chance for k -means to succeed.

We shall also explain the projection operation enforced in Step 2 of the algorithm. Given that each column of \mathbf{Y}^* has unit ℓ_2 norm, projecting each column of $\mathbf{U} \mathbf{U}^\top$ onto the unit sphere ensures that no column of \mathbf{Y} has an abnormal size. Note, however, that this projection step is non-essential and is introduced here primarily to simplify the mathematical analysis. Spectral clustering is expected to succeed even in the absence of such a projection step (Löffler *et al.*, 2021).

Remark 3.7. Another variation of spectral clustering is to directly apply the k -means algorithm to cluster the rows of \mathbf{U} (or some properly rescaled version of them) (Löffler *et al.*, 2021). To explain the rationale, we note that under the assumption (3.39), \mathbf{U}^* necessarily consists of r blocks of identical rows as follows:

$$\mathbf{U}^* = \sqrt{\frac{r}{n}} \begin{pmatrix} \mathbf{1}_{n/r} \boldsymbol{\nu}_1^\top \\ \vdots \\ \mathbf{1}_{n/r} \boldsymbol{\nu}_r^\top \end{pmatrix},$$

where $\boldsymbol{\nu}_1^\top, \dots, \boldsymbol{\nu}_r^\top$ are the orthonormal rows of the matrix \mathbf{U}_θ (cf. (3.44)). Consequently, clustering the rows of \mathbf{U}^* reveals exactly the true cluster assignments of all individuals. The idea of our spectral analysis below applies to this method as well; we leave it to the reader as an exercise.

3.5.3 Performance guarantees

Now, we turn to characterizing the clustering performance of the above spectral algorithm. We shall focus attention on the mis-clustering rate as the performance metric. As the cluster labels in $[r]$ can be arbitrarily permuted, the mis-clustering rate associated with the labels $\{\hat{\xi}_i\}$ returned by our algorithm is defined as

$$\ell_{\text{mis}}(\{\hat{\xi}_i\}, \{\xi_i^*\}) := \min_{\phi \in \Pi} \frac{1}{n} \sum_{i=1}^n \mathbb{1} \left\{ \phi(\hat{\xi}_i) \neq \xi_i^* \right\},$$

where Π is the set of permutations of $[r]$. In words, this metric captures the average number of mislabeled data points, after accounting for global permutation. For notational convenience, we shall set $\mathbf{M}^* := \mathbb{E}[\mathbf{X}^\top \mathbf{X}]$ and $\mathbf{E} := \mathbf{X}^\top \mathbf{X} - \mathbb{E}[\mathbf{X}^\top \mathbf{X}]$ throughout this section.

The first step towards analyzing the statistical accuracy of the spectral algorithm lies in developing a perturbation bound on $\|\mathbf{U}\mathbf{U}^\top - \mathbf{U}^*\mathbf{U}^{*\top}\|$, where \mathbf{U}^* (cf. (3.44)) represents the leading rank- r eigenspace of \mathbf{M}^* . This can be accomplished via the Davis-Kahan theorem, which requires us to first control the size of the perturbation \mathbf{E} .

Lemma 3.9. Consider the settings in Section 3.5.1, and suppose $p \gtrsim r \log^3(n + p)$. Then with probability at least $1 - O((n + p)^{-10})$, one has

$$\|\mathbf{E}\| \lesssim \frac{\Delta n \sqrt{\log(n + p)}}{\sqrt{r}} + \sqrt{np \log(n + p)} + n \log^2(n + p).$$

The proof of this lemma can be found in Section 3.5.4. Equipped with the above perturbation bound, we are ready to present our statistical guarantees for spectral clustering.

Theorem 3.10. Consider the setting and assumptions in Section 3.5.1, and suppose that $r = O(1)$ and $p \gtrsim \log^3 n$. With probability at least $1 - O(p^{-10})$, the mis-clustering rate of the spectral algorithm in Section 3.5.2 achieves

$$\ell_{\text{mis}}(\{\hat{\xi}_i\}, \{\xi_i^*\}) = o(1),$$

with the proviso that

$$\log(n + p) = o(\Delta) \quad \text{and} \quad \left(\frac{p \log(n + p)}{n} \right)^{1/4} = o(\Delta). \quad (3.46)$$

Before embarking on the proof of this theorem, we discuss briefly the implications of this theorem. In order to ensure a vanishingly small mis-clustering rate, it suffices for the center separation Δ to exceed

$$\Delta \gtrsim \begin{cases} \text{poly log}(n + p), & \text{if } p \leq n, \\ \left(\frac{p}{n}\right)^{1/4} \text{poly log}(n + p), & \text{if } p \geq n. \end{cases}$$

This separation condition matches the minimax lower bound up to some logarithmic term (Cai and Zhang, 2018; Ndaoud, 2018). In particular, in the high-dimensional case where $p \geq n$, the required separation condition changes fairly gracefully with the aspect ratio p/n .

Remark 3.8. As alluded to previously, Assumption 3.2 can be significantly relaxed. For example, the Gaussianity assumption therein is unnecessary; (almost) exact clustering is plausible once the minimum center separation exceeds a certain threshold, regardless of how $\{\theta_i^*\}$ are generated. To achieve this generality, however, the algorithm might need to be properly modified. Roughly speaking, in addition to \mathbf{U} , it is sensible to also exploit information contained in the eigenvalues of $\mathbf{X}^\top \mathbf{X}$ (which is crucial for, say, the scenario where all centers $\{\theta_i^*\}$ are perfectly aligned except for the scaling factors). We recommend the readers to Löffler *et al.* (2021) for detailed discussions.

Proof of Theorem 3.10. The proof consists of two steps: controlling the perturbation $\|\mathbf{U}\mathbf{U}^\top - \mathbf{U}^*\mathbf{U}^{*\top}\|_F$ (and hence $\|\mathbf{Y} - \mathbf{Y}^*\|_F$), and demonstrating that the follow-up k -means performs well.

The first step is to control $\|\mathbf{U}\mathbf{U}^\top - \mathbf{U}^*\mathbf{U}^{*\top}\|_F$, built upon Lemma 3.9 and a lower bound on the spectral gap of \mathbf{M}^* . Observe that

$$\lambda_r(\mathbf{M}^*) - \lambda_{r+1}(\mathbf{M}^*) = \lambda_{\min}(\mathbf{F}^* \boldsymbol{\Theta}^{*\top} \boldsymbol{\Theta}^* \mathbf{F}^{*\top}) = \frac{n}{r} \lambda_{\min}(\boldsymbol{\Theta}^{*\top} \boldsymbol{\Theta}^*),$$

where the last identity holds since $\mathbf{F}^* \mathbf{F}^{*\top} = \frac{n}{r} \mathbf{I}_r$ according to the definition (3.41). This motivates us to look at the spectral property of $\boldsymbol{\Theta}^{*\top} \boldsymbol{\Theta}^*$. Given that $\boldsymbol{\Theta}^*$ is composed of i.i.d. Gaussian entries (cf. Assumption 3.2), invoking the bound (3.28c) with proper rescaling gives

$$\|\boldsymbol{\Theta}^{*\top} \boldsymbol{\Theta}^* - \mathbb{E}[\boldsymbol{\Theta}^{*\top} \boldsymbol{\Theta}^*]\| \lesssim \Delta^2 \left(\sqrt{\frac{r \log p}{p}} + \frac{r \log^2 p}{p} \right) \leq \frac{\Delta^2}{4}$$

with probability exceeding $1 - O(p^{-10})$, provided that $p \geq C_2 r \log^2 p$ for some sufficiently large constant $C_2 > 0$. Further, it is self-evident that $\mathbb{E}[\boldsymbol{\Theta}^{*\top} \boldsymbol{\Theta}^*] = \frac{1}{2} \Delta^2 \mathbf{I}_r$. Weyl's inequality (see Lemma 2.2) then guarantees that

$$\begin{aligned} \lambda_{\min}(\boldsymbol{\Theta}^{*\top} \boldsymbol{\Theta}^*) &\geq \lambda_{\min}(\mathbb{E}[\boldsymbol{\Theta}^{*\top} \boldsymbol{\Theta}^*]) - \|\boldsymbol{\Theta}^{*\top} \boldsymbol{\Theta}^* - \mathbb{E}[\boldsymbol{\Theta}^{*\top} \boldsymbol{\Theta}^*]\| \\ &\geq \Delta^2/2 - \Delta^2/4 = \Delta^2/4. \end{aligned}$$

Combine the preceding inequalities to arrive at

$$\lambda_r(\mathbf{M}^*) - \lambda_{r+1}(\mathbf{M}^*) = \frac{n}{r} \lambda_{\min}(\boldsymbol{\Theta}^{*\top} \boldsymbol{\Theta}^*) \geq \frac{n \Delta^2}{4r}. \quad (3.47)$$

By virtue of Lemma 3.9 and (3.47), if the following condition

$$\Delta \geq C_1 \max \left\{ \left(\frac{r^2 p \log(n+p)}{n} \right)^{1/4}, \sqrt{r} \log(n+p) \right\}$$

holds for some large enough constant $C_1 > 0$, then it is guaranteed that $\|\mathbf{E}\| \leq (1 - 1/\sqrt{2})(\lambda_r(\mathbf{M}^*) - \lambda_{r+1}(\mathbf{M}^*))$. This in turn allows us to invoke the Davis-Kahan theorem (namely, Corollary 2.8) to obtain

$$\|\mathbf{U}\mathbf{U}^\top - \mathbf{U}^*\mathbf{U}^{*\top}\|_F \leq \frac{\sqrt{2r} \|\mathbf{E}\|}{\lambda_r(\mathbf{M}^*) - \lambda_{r+1}(\mathbf{M}^*)} \leq \sqrt{r} \varepsilon \quad (3.48)$$

with probability exceeding $1 - O(p^{-8})$, where the last line arises from (3.47) and Lemma 3.9, and

$$\varepsilon \asymp \frac{\Delta \sqrt{r \log(n+p)} + r \sqrt{\frac{p \log(n+p)}{n}} + r \log^2(n+p)}{\Delta^2}.$$

From the construction of \mathbf{Y} and (3.45), one can propagate the bound (3.48) to $\|\mathbf{Y} - \mathbf{Y}^*\|_{\text{F}}$ as follows:

$$\begin{aligned} \|\mathbf{Y} - \mathbf{Y}^*\|_{\text{F}}^2 &\stackrel{\text{(i)}}{=} \left\| \mathcal{P} \left(\sqrt{\frac{n}{r}} \mathbf{U} \mathbf{U}^\top \right) - \mathbf{Y}^* \right\|_{\text{F}}^2 \stackrel{\text{(ii)}}{\leq} 4 \left\| \sqrt{\frac{n}{r}} \mathbf{U} \mathbf{U}^\top - \mathbf{Y}^* \right\|_{\text{F}}^2 \\ &= \frac{4n}{r} \left\| \mathbf{U} \mathbf{U}^\top - \mathbf{U}^* \mathbf{U}^{*\top} \right\|_{\text{F}}^2 \lesssim \varepsilon^2 n. \end{aligned} \quad (3.49)$$

Here, the first identity (i) holds since the operator \mathcal{P} is invariant to global scaling. Regarding the inequality (ii), it follows from standard inequality regarding Euclidean projection (e.g., Soltanolkotabi (2019, Lemma 15)), which we postpone to the end of this proof.

The next step then amounts to translating the perturbation bound (3.49) into clustering accuracy guarantees (after k -means is applied). This is accomplished through the following key lemma, to be established in Section 3.5.4.

Lemma 3.11. Suppose that the matrix \mathbf{Y} obtained in the spectral algorithm in Section 3.5.2 satisfies

$$\|\mathbf{Y} - \mathbf{Y}^*\|_{\text{F}}^2 \leq \varepsilon^2 n, \quad (3.50)$$

where $\varepsilon > 0$ is a quantity obeying $\varepsilon \leq c_3 r^{-4}$ for some sufficiently small constant $c_3 > 0$. Then the mis-clustering rate obeys

$$\ell_{\text{mis}}(\{\hat{\xi}_i\}, \{\xi_i^*\}) \leq 2r\varepsilon^{1/4}.$$

As a consequence of Lemma 3.11, the mis-clustering rate is $o(1)$ as long as $\varepsilon r^4 = o(1)$, a condition that is guaranteed under the assumptions (3.46) and $r = O(1)$. This establishes Theorem 3.10.

Proof of the inequality (ii) in (3.49). For any vector \mathbf{v} residing in the unit sphere and any other vector \mathbf{w} , we have

$$\|\mathcal{P}(\mathbf{w}) - \mathbf{w}\|_2^2 = \|\mathcal{P}(\mathbf{w}) - \mathbf{v} + \mathbf{v} - \mathbf{w}\|_2^2$$

$$\begin{aligned}
&= \|\mathcal{P}(\mathbf{w}) - \mathbf{v}\|_2^2 + \|\mathbf{v} - \mathbf{w}\|_2^2 + 2\langle \mathcal{P}(\mathbf{w}) - \mathbf{v}, \mathbf{v} - \mathbf{w} \rangle \\
&\geq \|\mathcal{P}(\mathbf{w}) - \mathbf{v}\|_2^2 + \|\mathcal{P}(\mathbf{w}) - \mathbf{w}\|_2^2 + 2\langle \mathcal{P}(\mathbf{w}) - \mathbf{v}, \mathbf{v} - \mathbf{w} \rangle,
\end{aligned}$$

where the last inequality follows since \mathcal{P} denotes the projection onto the unit sphere and \mathbf{v} lies in the unit sphere. Cancelling out the common term $\|\mathcal{P}(\mathbf{w}) - \mathbf{w}\|_2^2$ and invoking Cauchy-Schwarz lead to

$$\|\mathcal{P}(\mathbf{w}) - \mathbf{v}\|_2^2 \leq -2\langle \mathcal{P}(\mathbf{w}) - \mathbf{v}, \mathbf{v} - \mathbf{w} \rangle \leq 2\|\mathcal{P}(\mathbf{w}) - \mathbf{v}\|_2\|\mathbf{v} - \mathbf{w}\|_2,$$

and therefore,

$$\|\mathcal{P}(\mathbf{w}) - \mathbf{v}\|_2 \leq 2\|\mathbf{w} - \mathbf{v}\|_2, \quad \text{or} \quad \|\mathcal{P}(\mathbf{w}) - \mathbf{v}\|_2^2 \leq 4\|\mathbf{w} - \mathbf{v}\|_2^2.$$

This inequality clearly extends to the matrix counterpart, thus establishing the claimed result. \square

3.5.4 Proof of auxiliary lemmas

Proof of Lemma 3.9. To begin with, let us decompose \mathbf{E} as follows

$$\begin{aligned}
\mathbf{E} &= \mathbf{X}^\top \mathbf{X} - \mathbb{E}[\mathbf{X}^\top \mathbf{X}] = (\Theta^* \mathbf{F}^{*\top} + \mathbf{Z})^\top (\Theta^* \mathbf{F}^{*\top} + \mathbf{Z}) - \mathbb{E}[\mathbf{X}^\top \mathbf{X}] \\
&= \mathbf{F}^* \Theta^{*\top} \mathbf{Z} + \mathbf{Z}^\top \Theta^* \mathbf{F}^{*\top} + \mathbf{Z}^\top \mathbf{Z} - p\mathbf{I}_n,
\end{aligned} \tag{3.51}$$

where we have used the notation in (3.40) and (3.41), as well as the identity (3.42). As it turns out, similar terms have already been controlled in the proof for PCA (see Section 3.3.4). More precisely, the first term in (3.51) obeys

$$\begin{aligned}
\|\mathbf{F}^* \Theta^{*\top} \mathbf{Z}\| &\stackrel{\text{(i)}}{=} \sqrt{\frac{n}{r}} \|\Theta^{*\top} \mathbf{Z}\| \stackrel{\text{(ii)}}{\lesssim} \sqrt{\frac{n}{r}} \cdot \frac{\Delta}{\sqrt{p}} \sqrt{np \log(n+p)} \\
&\asymp \frac{\Delta n \sqrt{\log(n+p)}}{\sqrt{r}}
\end{aligned}$$

with probability exceeding $1 - O((n+p)^{-10})$, provided that $p \gtrsim r \log^3(n+p)$. Here, the first relation (i) holds true since $\sqrt{\frac{r}{n}} \mathbf{F}^*$ contains orthonormal columns and the spectral norm is unitarily invariant, while (ii) invokes the high-probability bound (3.28a). When it comes to the third term of (3.51), the bound (3.28c) readily implies that

$$\|\mathbf{Z}^\top \mathbf{Z} - p\mathbf{I}_n\| \lesssim \sqrt{np \log(n+p)} + n \log^2(n+p)$$

with probability at least $1 - O((n + p)^{-10})$. Substituting the preceding two bounds into (3.51) and applying the triangle inequality, we reach

$$\begin{aligned}\|\mathbf{E}\| &\leq 2\|\mathbf{F}^* \boldsymbol{\Theta}^{*\top} \mathbf{Z}\| + \|\mathbf{Z}^\top \mathbf{Z} - p\mathbf{I}_n\| \\ &\lesssim \frac{\Delta n \sqrt{\log(n + p)}}{\sqrt{r}} + \sqrt{np \log(n + p)} + n \log^2(n + p).\end{aligned}$$

Proof of Lemma 3.11 (analysis for k -means). Given the class labels $\{\xi_i\}_{i=1}^n$, the optimization of the cluster centers $\{\boldsymbol{\vartheta}_i\}_{i=1}^r$ in the k -means formulation (3.43) is achieved by the sample means of each cluster. Thus, the k -means formulation (3.43) can be equivalently posed as solving

$$\underset{\mathcal{C} \in \Xi}{\text{minimize}} \sum_{l=1}^r \sum_{i \in \mathcal{C}_l} \left\| \mathbf{y}_i - \frac{1}{|\mathcal{C}_l|} \sum_{j \in \mathcal{C}_l} \mathbf{y}_j \right\|_2^2 \quad (3.52)$$

where $\mathcal{C} = \{\mathcal{C}_1, \dots, \mathcal{C}_r\}$ represents the cluster assignment, and Ξ denotes the set of all r -partitions of $[n]$ (i.e., r disjoint subsets whose union equals $[n]$).

In order to tackle this formulation, a key ingredient of the proof lies in the following deviation bound that allows one to replace \mathbf{y}_i with the truth \mathbf{y}_i^* , as long as the cluster size is sufficiently large.

Claim 3.1. Consider any set $\mathcal{S} \subseteq [n]$ with cardinality $c_s n$ for some quantity $c_s > 0$. Suppose that (3.50) holds with $\varepsilon \leq c_s^2$. Then one has

$$\left| \sum_{i \in \mathcal{S}} \left\| \mathbf{y}_i - \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j \right\|_2^2 - \sum_{i \in \mathcal{S}} \left\| \mathbf{y}_i^* - \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j^* \right\|_2^2 \right| \leq 6c_s \sqrt{\varepsilon n}. \quad (3.53)$$

With Claim 3.1 in place, we are positioned to establish Lemma 3.11 by contradiction; that is, we intend to demonstrate that any cluster assignment that differs too much from the ground-truth clusters cannot possibly be the k -means solution. In what follows, we denote by \mathcal{C}_l^* the l -th ground-truth cluster ($1 \leq l \leq r$), and let $\{\mathcal{C}_1, \dots, \mathcal{C}_r\}$ represent the minimizer of (3.52) whenever it is clear from the context.

Step 1: developing an upper bound on (3.52). To begin with, we derive an upper bound on the optimal objective value of (3.52), which serves as

a reference in assessing the (sub)-optimality of other cluster assignments. By virtue of Claim 3.1 and the assumption $|\mathcal{C}_l^*| = n/r$, one has

$$\left| \sum_{i \in \mathcal{C}_l^*} \left\| \mathbf{y}_i - \frac{1}{|\mathcal{C}_l^*|} \sum_{j \in \mathcal{C}_l^*} \mathbf{y}_j \right\|_2^2 - \sum_{i \in \mathcal{C}_l^*} \left\| \mathbf{y}_i^* - \frac{1}{|\mathcal{C}_l^*|} \sum_{j \in \mathcal{C}_l^*} \mathbf{y}_j^* \right\|_2^2 \right| \leq \frac{6\sqrt{\varepsilon}n}{r},$$

with the proviso that $\varepsilon \leq 1/r^2$. Note that by construction, for the “ideal” fitting, one has $\sum_{i \in \mathcal{C}_l^*} \left\| \mathbf{y}_i^* - \frac{1}{|\mathcal{C}_l^*|} \sum_{j \in \mathcal{C}_l^*} \mathbf{y}_j^* \right\|_2^2 = 0$. Using this fact and summing the above inequality over all $1 \leq l \leq r$, we arrive at

$$\sum_{l=1}^r \sum_{i \in \mathcal{C}_l^*} \left\| \mathbf{y}_i - \frac{1}{|\mathcal{C}_l^*|} \sum_{j \in \mathcal{C}_l^*} \mathbf{y}_j \right\|_2^2 \leq 6\sqrt{\varepsilon}n. \quad (3.54)$$

Consequently, due to the assumed optimality of \mathcal{C} w.r.t. (3.52), replacing $\{\mathcal{C}_l^*\}_{l=1}^r$ in (3.54) by $\{\mathcal{C}_l\}_{l=1}^r$ can only further improve the objective value:

$$\begin{aligned} \sum_{l=1}^r \sum_{i \in \mathcal{C}_l} \left\| \mathbf{y}_i - \frac{1}{|\mathcal{C}_l|} \sum_{j \in \mathcal{C}_l} \mathbf{y}_j \right\|_2^2 &\leq \sum_{l=1}^r \sum_{i \in \mathcal{C}_l^*} \left\| \mathbf{y}_i - \frac{1}{|\mathcal{C}_l^*|} \sum_{j \in \mathcal{C}_l^*} \mathbf{y}_j \right\|_2^2 \\ &\leq 6\sqrt{\varepsilon}n. \end{aligned} \quad (3.55)$$

Step 2: showing that no cluster can be too large. Suppose that there exists a cluster \mathcal{C}_l ($1 \leq l \leq r$) that is too large in the sense that

$$|\mathcal{C}_l| \geq \frac{(1 + c_\varepsilon)n}{r} \quad (3.56)$$

for some quantity $c_\varepsilon > 0$. We would like to show that this is impossible unless c_ε is small; in fact, in light of Claim 3.1, we need only to establish a lower bound on the second term in (3.53) (with $\mathcal{S} = \mathcal{C}_l$) so that it leads to a contradiction with the upper bound (3.55). Towards this end, we start with the elementary decomposition of the sum of squared errors:

$$\begin{aligned} \sum_{i \in \mathcal{C}_l} \left\| \mathbf{y}_i^* - \frac{1}{|\mathcal{C}_l|} \sum_{j \in \mathcal{C}_l} \mathbf{y}_j^* \right\|_2^2 &= \left(\sum_{i \in \mathcal{C}_l} \left\| \mathbf{y}_i^* \right\|_2^2 \right) - |\mathcal{C}_l| \cdot \left\| \frac{1}{|\mathcal{C}_l|} \sum_{j \in \mathcal{C}_l} \mathbf{y}_j^* \right\|_2^2 \\ &= |\mathcal{C}_l| \left(1 - \left\| \frac{1}{|\mathcal{C}_l|} \sum_{j \in \mathcal{C}_l} \mathbf{y}_j^* \right\|_2^2 \right), \end{aligned} \quad (3.57)$$

where we have invoked the fact that $\|\mathbf{y}_i^*\|_2 = 1$. It thus comes down to controlling $\|\sum_{j \in \mathcal{C}_l} \mathbf{y}_j^*\|_2^2$. For notational convenience, for any $1 \leq \tau \leq r$, we set $n_{l,\tau} := |\mathcal{C}_l \cap \mathcal{C}_\tau^*|$ (namely, the number of points in \mathcal{C}_l coming from the τ -th ground-truth cluster), and let $\mathbf{y}_{(\tau)}^*$ represent the vector associated with the τ -th cluster (namely, $\mathbf{y}_{(\tau)}^* = \mathbf{y}_j^*$ for any $j \in \mathcal{C}_\tau^*$). Armed with this set of notation, we can write

$$\left\| \frac{1}{|\mathcal{C}_l|} \sum_{j \in \mathcal{C}_l} \mathbf{y}_j^* \right\|_2^2 = \left\| \sum_{\tau=1}^r \frac{n_{l,\tau}}{|\mathcal{C}_l|} \mathbf{y}_{(\tau)}^* \right\|_2^2 = \sum_{\tau=1}^r \left\| \frac{n_{l,\tau}}{|\mathcal{C}_l|} \mathbf{y}_{(\tau)}^* \right\|_2^2 = \frac{\sum_{\tau=1}^r n_{l,\tau}^2}{|\mathcal{C}_l|^2}. \quad (3.58)$$

Here, the penultimate identity holds since $\langle \mathbf{y}_{(i)}^*, \mathbf{y}_{(j)}^* \rangle = 0$ for any $i \neq j$, while the last relation relies on the fact that $\|\mathbf{y}_{(\tau)}^*\|_2 = 1$. Using $0 \leq n_{l,\tau} \leq n/r$ and $\sum_{\tau=1}^r n_{l,\tau} = |\mathcal{C}_l|$, we have

$$(3.58) \leq \frac{(\max_\tau n_{l,\tau})(\sum_\tau n_{l,\tau})}{|\mathcal{C}_l|^2} \leq \frac{(n/r) \cdot |\mathcal{C}_l|}{|\mathcal{C}_l|^2} \leq \frac{1}{1 + c_\varepsilon},$$

where the first inequality comes from the basic fact that $\|\mathbf{a}\|_2^2 \leq \|\mathbf{a}\|_\infty \|\mathbf{a}\|_1$ for any vector \mathbf{a} , and the last relation arises from the assumed cardinality constraint on \mathcal{C}_l (cf. (3.56)).

Substitution into (3.57) yields

$$\sum_{i \in \mathcal{C}_l} \left\| \mathbf{y}_i^* - \frac{1}{|\mathcal{C}_l|} \sum_{j \in \mathcal{C}_l} \mathbf{y}_j^* \right\|_2^2 \geq |\mathcal{C}_l| \left(1 - \frac{1}{1 + c_\varepsilon} \right) \geq c_\varepsilon \frac{n}{r}, \quad (3.59)$$

where the last inequality again arises from the assumption (3.56). This in turn demonstrates that

$$\begin{aligned} \sum_{l=1}^r \sum_{i \in \mathcal{C}_l} \left\| \mathbf{y}_i - \frac{1}{|\mathcal{C}_l|} \sum_{j \in \mathcal{C}_l} \mathbf{y}_j \right\|_2^2 &\geq \sum_{i \in \mathcal{C}_l} \left\| \mathbf{y}_i - \frac{1}{|\mathcal{C}_l|} \sum_{j \in \mathcal{C}_l} \mathbf{y}_j \right\|_2^2 \\ &\geq \sum_{i \in \mathcal{C}_l} \left\| \mathbf{y}_i^* - \frac{1}{|\mathcal{C}_l|} \sum_{j \in \mathcal{C}_l} \mathbf{y}_j^* \right\|_2^2 - 6 \frac{|\mathcal{C}_l|}{n} \sqrt{\varepsilon n} \\ &\geq c_\varepsilon \frac{n}{r} - 6\sqrt{\varepsilon n} > 6\sqrt{\varepsilon n}, \end{aligned} \quad (3.60)$$

where (3.60) results from Claim 3.1 when $\varepsilon \leq 1/r^2$, and the last inequality follows as long as $12\sqrt{\varepsilon} < c_\varepsilon/r$. Comparing this with (3.55) leads to contradiction with the optimality assumption of $\{\mathcal{C}_l\}$.

Step 3: showing that no cluster can be too small. Suppose now that there exists a cluster \mathcal{C}_i ($1 \leq i \leq r$) obeying

$$|\mathcal{C}_i| \leq \frac{(1 - c_\varepsilon(r-1))n}{r}.$$

Then from the pigeonhole principle, one can find another cluster \mathcal{C}_l ($1 \leq l \leq r$) with cardinality exceeding

$$|\mathcal{C}_l| \geq \frac{(1 + c_\varepsilon)n}{r};$$

otherwise the total size obeys $\sum_{k=1}^r |\mathcal{C}_k| < \frac{(1 - c_\varepsilon(r-1))n}{r} + (r-1)\frac{(1 + c_\varepsilon)n}{r} \leq n$ and $\{\mathcal{C}_l\}$ is infeasible. The above condition on $|\mathcal{C}_l|$ coincides with the assumption (3.56) in Step 2, which, as a result of previous arguments, cannot possibly hold. To conclude, for all $1 \leq i \leq r$, one necessarily has

$$|\mathcal{C}_i| > \frac{(1 - c_\varepsilon(r-1))n}{r} \geq \frac{(1 - c_\varepsilon)r)n}{r}. \quad (3.61)$$

Step 4: showing that each \mathcal{C}_l is mainly composed of points from a true (and distinct) cluster. Suppose that there exists a cluster \mathcal{C}_l ($1 \leq l \leq r$) whose dominant component obeys

$$\max_{1 \leq \tau \leq r} n_{l,\tau} \leq (1 - 2rc_\varepsilon)n/r,$$

where we recall that $n_{l,\tau} = |\mathcal{C}_l \cap \mathcal{C}_\tau^*|$. Under this assumption, we have

$$(3.58) \leq \frac{(\max_\tau n_{l,\tau})(\sum_\tau n_{l,\tau})}{|\mathcal{C}_l|^2} \leq \frac{[(1 - 2rc_\varepsilon)n/r] \cdot |\mathcal{C}_l|}{|\mathcal{C}_l|^2} \leq \frac{1 - 2rc_\varepsilon}{1 - rc_\varepsilon},$$

where the last inequality relies on the lower bound (3.61). Substitution into (3.57) gives

$$\begin{aligned} \sum_{i \in \mathcal{C}_l} \left\| \mathbf{y}_i^* - \frac{1}{|\mathcal{C}_l|} \sum_{j \in \mathcal{C}_l} \mathbf{y}_j^* \right\|_2^2 &\geq |\mathcal{C}_l| \left(1 - \frac{1 - 2rc_\varepsilon}{1 - rc_\varepsilon}\right) \stackrel{(i)}{\geq} \frac{(1 - rc_\varepsilon)n}{r} \cdot \frac{rc_\varepsilon}{1 - rc_\varepsilon} \\ &= c_\varepsilon n > 12\sqrt{\varepsilon}n, \end{aligned}$$

where (i) arises again from (3.61), and the last relation is valid once $c_\varepsilon > 12\sqrt{\varepsilon}$. This taken collectively with the inequality (3.60) yields

$$\sum_{l=1}^r \sum_{i \in \mathcal{C}_l} \left\| \mathbf{y}_i - \frac{1}{|\mathcal{C}_l|} \sum_{j \in \mathcal{C}_l} \mathbf{y}_j \right\|_2^2 \geq \sum_{i \in \mathcal{C}_l} \left\| \mathbf{y}_i^* - \frac{1}{|\mathcal{C}_l|} \sum_{j \in \mathcal{C}_l} \mathbf{y}_j^* \right\|_2^2 - 6\sqrt{\varepsilon}n$$

$$> 6\sqrt{\varepsilon}n,$$

which, however, contradicts the upper bound (3.55). Consequently, the dominant component in every cluster $1 \leq l \leq r$ must obey

$$\max_{1 \leq \tau \leq r} n_{l,\tau} > (1 - 2rc_\varepsilon)n/r. \quad (3.62)$$

In particular, if $2rc_\varepsilon < 1/2$, then $\max_{1 \leq \tau \leq r} n_{l,\tau} > n/(2r)$. An immediate consequence is that: the dominant components of the clusters $\{\mathcal{C}_l\}$ must come from distinct ground-truth clusters.

Step 5: putting all this together. Armed with the preceding bound (3.62) and the remark thereafter, it is straightforward to verify the following result on the mis-clustering rate:

$$\ell_{\text{mis}}(\{\widehat{\xi}_i\}, \{\xi_i^*\}) \leq 1 - \frac{\sum_{l=1}^r \max_{1 \leq \tau \leq r} n_{l,\tau}}{n} \leq 2rc_\varepsilon.$$

Finally, setting $c_\varepsilon = \varepsilon^{1/4}$ leads to the advertised result, provided that $\varepsilon \leq c_3r^{-4}$ for some sufficiently small constant $c_3 > 0$.

Proof of Claim 3.1. Before proceeding, let us take a quick look at how many columns of \mathbf{Y} might deviate considerably from their counterparts in \mathbf{Y}^* . To be precise, let us introduce the following set

$$\mathcal{N}_{\text{large}} := \left\{ i \mid \|\mathbf{y}_i - \mathbf{y}_i^*\|_2^2 \geq \varepsilon \right\}. \quad (3.63)$$

Clearly, its cardinality is necessarily bounded above by

$$|\mathcal{N}_{\text{large}}| \leq \frac{\|\mathbf{Y} - \mathbf{Y}^*\|_{\text{F}}^2}{\varepsilon} \leq \frac{\varepsilon^2 n}{\varepsilon} = \varepsilon n. \quad (3.64)$$

The starting point of the proof is the elementary identities

$$\begin{aligned} & \sum_{i \in \mathcal{S}} \left\| \mathbf{y}_i - \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j \right\|_2^2 - \sum_{i \in \mathcal{S}} \left\| \mathbf{y}_i^* - \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j^* \right\|_2^2 \\ &= \sum_{i \in \mathcal{S}} \|\mathbf{y}_i\|_2^2 - |\mathcal{S}| \cdot \left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j \right\|_2^2 - \left(\sum_{i \in \mathcal{S}} \|\mathbf{y}_i^*\|_2^2 - |\mathcal{S}| \cdot \left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j^* \right\|_2^2 \right) \\ &= |\mathcal{S}| \cdot \left(\left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j^* \right\|_2^2 - \left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j \right\|_2^2 \right), \end{aligned} \quad (3.65)$$

where the last inequality follows from the fact that $\|\mathbf{y}_i\|_2 = \|\mathbf{y}_i^*\|_2 = 1$. To bound the right-hand side of (3.65), we make the observation that

$$\begin{aligned} \left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} (\mathbf{y}_j - \mathbf{y}_j^*) \right\|_2 &\leq \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S} \setminus \mathcal{N}_{\text{large}}} \|\mathbf{y}_j - \mathbf{y}_j^*\|_2 + \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S} \cap \mathcal{N}_{\text{large}}} \|\mathbf{y}_j - \mathbf{y}_j^*\|_2 \\ &\stackrel{(i)}{\leq} \sqrt{\varepsilon} + \frac{2|\mathcal{N}_{\text{large}}|}{|\mathcal{S}|} \stackrel{(ii)}{\leq} \sqrt{\varepsilon} + \frac{2\varepsilon}{c_s} \leq 3\sqrt{\varepsilon}. \end{aligned} \quad (3.66)$$

Here, (i) holds true since any column outside $\mathcal{N}_{\text{large}}$ satisfies $\|\mathbf{y}_j - \mathbf{y}_j^*\|_2 < \sqrt{\varepsilon}$, and any column coming from $\mathcal{N}_{\text{large}}$ obeys $\|\mathbf{y}_j - \mathbf{y}_j^*\|_2 \leq \|\mathbf{y}_j\|_2 + \|\mathbf{y}_j^*\|_2 = 2$; (ii) follows from (3.64) and the assumption $|\mathcal{S}| = c_s n$; and the last inequality relies on the assumption $\varepsilon \leq c_s^2$. In addition,

$$\left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j^* \right\|_2 \leq \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \|\mathbf{y}_j^*\|_2 = 1; \quad (3.67a)$$

$$\left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j \right\|_2 \leq \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \|\mathbf{y}_j\|_2 = 1. \quad (3.67b)$$

Combining (3.66) and (3.67) and applying the triangle inequality, we arrive at

$$\begin{aligned} &\left| \left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j \right\|_2^2 - \left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j^* \right\|_2^2 \right| \\ &= \left| \left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j \right\|_2 - \left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j^* \right\|_2 \right| \cdot \left| \left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j \right\|_2 + \left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j^* \right\|_2 \right| \\ &\leq 2 \left| \left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j \right\|_2 - \left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j^* \right\|_2 \right| \\ &\leq 2 \left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} (\mathbf{y}_j - \mathbf{y}_j^*) \right\|_2 \leq 6\sqrt{\varepsilon}. \end{aligned} \quad (3.68)$$

Finally, plugging in (3.68) into (3.65), we conclude that

$$\begin{aligned} &\left| \sum_{i \in \mathcal{S}} \left\| \mathbf{y}_i - \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j \right\|_2^2 - \sum_{i \in \mathcal{S}} \left\| \mathbf{y}_i^* - \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j^* \right\|_2^2 \right| \\ &= |\mathcal{S}| \cdot \left| \left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j \right\|_2^2 - \left\| \frac{1}{|\mathcal{S}|} \sum_{j \in \mathcal{S}} \mathbf{y}_j^* \right\|_2^2 \right| \end{aligned}$$

$$\leq 6\sqrt{\varepsilon}|\mathcal{S}| = 6c_s\sqrt{\varepsilon}n$$

as claimed, where the last relation holds true since $|\mathcal{S}| = c_s n$.

3.6 Ranking from pairwise comparisons

The ranking task—which seeks to identify a consistent ordering of several items based on (partially) revealed preference information about them—is encountered in numerous contexts including web search, crowd sourcing, social choice, peer grading, and so on (Dwork *et al.*, 2001; Chen *et al.*, 2013; Caplin and Nalebuff, 1991; Shah *et al.*, 2013). Of particular interest is the “preference-based” observation model, in which we are only given relative comparisons of a few items (as opposed to individual scores of them). In practice, comparison data of this kind abound, partly because humans often find it easier to make a preference over two or a couple of items than to assign specific ratings to many individual ones. The emergence of crowdsourcing platforms such as Amazon Mechanical Turk further widens the availability of comparison data, where binary judgments over pairs of items are often solicited from a pool of non-experts. In this section, we concentrate on pairwise comparisons and explore the potential of spectral methods for the ranking task.

3.6.1 The Bradley-Terry-Luce model and assumptions

To formulate the problem in a statistically sound manner, we introduce a classical parametric model, called the Bradley-Terry-Luce (BTL) model (Bradley and Terry, 1952; Ford Jr, 1957; Luce, 2012), to describe the generating process of pairwise comparisons.

Latent preference scores. Imagine that there are n items to be ranked. A key component of the BTL model is the assignment of a latent preference score to each item; more concretely, the BTL model hypothesizes on the existence of an unseen preference score vector

$$\mathbf{w}^* = [w_1^*, w_2^*, \dots, w_n^*]^\top, \quad (3.69)$$

with $w_i^* > 0$ assigned to the i -th item ($1 \leq i \leq n$). The ranks of these items are therefore determined exclusively by their (relative) preference scores: an item with a larger score is ranked higher. Throughout this section, we denote by κ a sort of condition number as follows

$$\kappa := \frac{\max_{1 \leq i \leq n} w_i^*}{\min_{1 \leq i \leq n} w_i^*}. \quad (3.70)$$

Pairwise comparisons. Equipped with the aforementioned score vector, the BTL model posits that: the probability of an item winning a paired comparison is determined entirely by the relative scores of the two items involved. To be precise, when comparing every pair (i, j) of items, the model assumes that

$$\mathbb{P}\{\text{item } j \text{ is preferred over item } i\} = \frac{w_j^*}{w_i^* + w_j^*}, \quad (3.71)$$

asserting that an item assigned a higher preference score is more likely to win. In this section, we assume access to a comparison between every pair of items. To be precise, for each pair (i, j) ($1 \leq i < j \leq n$), we observe an independent binary comparison outcome $y_{i,j}$ following the BTL model (3.71):

$$y_{i,j} = \begin{cases} 1, & \text{with probability } \frac{w_j^*}{w_i^* + w_j^*}, \\ 0, & \text{otherwise,} \end{cases}$$

where $y_{i,j} = 1$ means item j beats item i and $y_{i,j} = 0$ otherwise. By convention, we set $y_{i,j} = 1 - y_{j,i}$ for all $i > j$.

Goal. With the BTL parametric model in mind, a natural strategy is to start by estimating the underlying scores $\{w_i^*\}$ based on the pairwise comparisons in hand, followed by a ranking step performed in accordance with the estimated scores. In this section, we shall focus on characterizing the statistical accuracy of spectral methods in accomplishing the meta task of preference score estimation, and will remark in passing on the ranking step that follows. Obviously, from (3.71), we can only hope for estimating $\{w_i^*\}$ up to some global scaling ambiguity.

3.6.2 A spectral ranking algorithm

At first glance, the recipe we have introduced for designing spectral methods seems to have no direct bearing on the BTL model. Somewhat unexpectedly, a closer inspection unveils an intimate connection between the BTL model and a reversible Markov chain, whose stationary distribution embodies crucial information about the score vector of interest. This in turn lays a solid foundation for the spectral algorithm described below, originally developed by Negahban *et al.* (2016).

The first step is to convert the pairwise comparison data $\{y_{i,j}\}_{i \neq j}$ into a probability transition matrix $\mathbf{P} = [P_{i,j}]_{1 \leq i,j \leq n}$, in a way that

$$P_{i,j} = \begin{cases} \frac{1}{n} y_{i,j}, & \text{if } i \neq j, \\ 1 - \sum_{j:j \neq i} \frac{1}{n} y_{i,j}, & \text{otherwise.} \end{cases} \quad (3.72)$$

By construction of \mathbf{P} , all of its entries are non-negative and the entries in each row sum up to one, thus confirming that \mathbf{P} is a probability transition matrix. The spectral algorithm then computes the leading left eigenvector $\boldsymbol{\pi}$ of \mathbf{P} , returning it as the estimate for the underlying score vector \mathbf{w}^* .

To make sense of the rationale behind this algorithm, it is helpful to look at the mean $\mathbf{P}^* = [P_{i,j}^*]_{1 \leq i,j \leq n} := \mathbb{E}[\mathbf{P}]$, which obeys

$$P_{i,j}^* = \begin{cases} \frac{1}{n} \frac{w_j^*}{w_i^* + w_j^*}, & \text{if } i \neq j, \\ 1 - \frac{1}{n} \sum_{j:j \neq i} \frac{w_j^*}{w_i^* + w_j^*}, & \text{otherwise.} \end{cases} \quad (3.73)$$

Clearly, this matrix \mathbf{P}^* is a probability transition matrix as well. As can be straightforwardly verified, the vector $\boldsymbol{\pi}^* = [\pi_i^*]_{1 \leq i \leq n}$ defined by

$$\boldsymbol{\pi}^* = \frac{1}{\mathbf{1}^\top \mathbf{w}^*} \mathbf{w}^* \quad (3.74)$$

satisfies the following conditions:

- $\boldsymbol{\pi}^*$ is a probability vector (i.e., $\pi_i^* \geq 0$ for all i and $\sum_i \pi_i^* = 1$);
- $\boldsymbol{\pi}^*$ satisfies the detailed balance equations as follows:

$$\pi_i^* P_{i,j}^* = \pi_j^* P_{j,i}^*, \quad \text{for all } (i, j).$$

Classical Markov chain theory (Brémaud, 2013) thus tells us that \mathbf{P}^* represents a reversible Markov chain, whose stationary distribution is precisely given by $\boldsymbol{\pi}^*$ (this can easily be verified using the definition of the stationary distribution) and corresponds to the normalized preference scores. As a consequence, we hold the intuition that: $\boldsymbol{\pi}$ is close to $\boldsymbol{\pi}^*$ —and hence \mathbf{w}^* up to some global scaling—as long as \mathbf{P} approximates \mathbf{P}^* reasonably well.

3.6.3 Performance guarantees

This subsection develops theoretical support for the above spectral ranking algorithm, based on the eigenvector perturbation theory developed previously for probability transition matrices in Section 2.5. For notational convenience, we shall use $\mathbf{E} := \mathbf{P} - \mathbf{P}^*$ to denote the difference of the above two transition matrices of interest.

By virtue of Theorem 2.10, the perturbation of the stationary distribution of a reversible Markov chain \mathbf{P}^* is dictated by two important quantities: (i) the spectral gap $1 - \max\{\lambda_2(\mathbf{P}^*), -\lambda_n(\mathbf{P}^*)\}$, and (ii) the noise size $\|\mathbf{E}\|_{\boldsymbol{\pi}^*}$ (recall the definition of $\|\cdot\|_{\boldsymbol{\pi}^*}$ in Section 2.5.1). These two quantities are controlled respectively via the following two lemmas, whose proofs can be found in Section 3.6.4.

Lemma 3.12. Consider the settings and notation in Sections 3.6.1 and 3.6.2. It follows that

$$1 - \max\{\lambda_2(\mathbf{P}^*), -\lambda_n(\mathbf{P}^*)\} \geq \frac{1}{2\kappa^2},$$

where we recall the definition of κ in (3.70).

Lemma 3.13. Consider the settings and notation in Sections 3.6.1 and 3.6.2, and recall that $\mathbf{E} := \mathbf{P} - \mathbf{P}^*$. With probability at least $1 - O(n^{-8})$,

$$\|\mathbf{E}\|_{\boldsymbol{\pi}^*} \leq \sqrt{\kappa} \|\mathbf{E}\| \lesssim \sqrt{\frac{\kappa \log n}{n}}.$$

Now we are well prepared to assess the quality of the spectral estimate, as summarized below, whose proof is given at the end of this subsection.

Theorem 3.14. Consider the settings and algorithm in Sections 3.6.1 and 3.6.2. Suppose that $n \geq C\kappa^5 \log n$ for some sufficiently large constant $C > 0$. Then with probability exceeding $1 - O(n^{-8})$, one has

$$\frac{\|\boldsymbol{\pi} - \boldsymbol{\pi}^*\|_2}{\|\boldsymbol{\pi}^*\|_2} \lesssim \kappa^{2.5} \sqrt{\frac{\log n}{n}}. \quad (3.75)$$

Given the construction (3.74) of $\boldsymbol{\pi}^*$, this theorem implies the existence of a scalar $z > 0$ such that

$$\frac{\|z\boldsymbol{\pi} - \mathbf{w}^*\|_2}{\|\mathbf{w}^*\|_2} \lesssim \kappa^{2.5} \sqrt{\frac{\log n}{n}}$$

holds with high probability. It is worth noting that one cannot possibly retrieve the global scaling factor z , due to the invariance of the BTL observation model under global scaling (cf. (3.71)).

To interpret the effectiveness of this theorem, consider, for example, the case when $\kappa = O(1)$ (so that all the latent scores w_i^* are about the same order). Theorem 3.14 tells us that the relative estimation error of $\boldsymbol{\pi}$ is vanishing as the number n of items increases. As it turns out, this statistical error rate (3.75) is near minimax-optimal up to a logarithmic factor; see Negahban *et al.* (2016, Theorem 3). In fact, with a more careful analysis, one can further eliminate this extra $\log n$ factor and establish (orderwise) minimax optimality of this algorithm, as has been done in Chen *et al.* (2019b, Theorem 5.2).

Caution needs to be exercised, however, that high score estimation accuracy alone does not necessarily imply appealing ranking accuracy. For instance, if the goal is to identify the top- K ranked items—a problem commonly referred to as “top- K ranking” (Chen and Suh, 2015)—then the ranking accuracy also relies heavily on the separation between the score of the K -th ranked item and that of the $(K + 1)$ -th ranked item (namely, whether the set of top- K ranked items is sufficiently distinguishable from the remaining ones). Fortunately, the spectral ranking algorithm introduced in this section remains minimax optimal when it comes to top- K ranking, through a refined ℓ_∞ perturbation theory to be introduced in Chapter 4. The interested reader is referred to Chen *et al.* (2019b) for details.

Proof of Theorem 3.14. Invoke Theorem 2.10 to see that

$$\begin{aligned}\|\boldsymbol{\pi} - \boldsymbol{\pi}^*\|_{\boldsymbol{\pi}^*} &\leq \frac{\|\boldsymbol{\pi}^{*\top} \mathbf{E}\|_{\boldsymbol{\pi}^*}}{1 - \max\{\lambda_2(\mathbf{P}^*), -\lambda_n(\mathbf{P}^*)\} - \|\mathbf{E}\|_{\boldsymbol{\pi}^*}} \\ &\leq 4\kappa^2 \|\boldsymbol{\pi}^{*\top} \mathbf{E}\|_{\boldsymbol{\pi}^*},\end{aligned}\tag{3.76}$$

provided that

$$1 - \max\{\lambda_2(\mathbf{P}^*), -\lambda_n(\mathbf{P}^*)\} - \|\mathbf{E}\|_{\boldsymbol{\pi}^*} \geq 1/(4\kappa^2).\tag{3.77}$$

From Lemma 3.12 and Lemma 3.13, we know that Condition (3.77) holds true with probability at least $1 - O(n^{-8})$, with the proviso that $n \geq C\kappa^5 \log n$ for some sufficiently large constant $C > 0$.

Additionally, letting $\pi_{\min}^* := \min_i \pi_i^*$ and $\pi_{\max}^* := \max_i \pi_i^*$, we can easily see from the definition of $\|\cdot\|_{\boldsymbol{\pi}^*}$ (i.e., $\|\mathbf{v}\|_{\boldsymbol{\pi}^*} = \sqrt{\sum_i \pi_i^* v_i^2}$ for any vector \mathbf{v}) that

$$\|\mathbf{v}\|_{\boldsymbol{\pi}^*} \stackrel{(i)}{\leq} \sqrt{\pi_{\max}^*} \|\mathbf{v}\|_2, \quad \text{and} \quad \|\mathbf{v}\|_2 \stackrel{(ii)}{\leq} \frac{1}{\sqrt{\pi_{\min}^*}} \|\mathbf{v}\|_{\boldsymbol{\pi}^*},$$

which allows us to further obtain

$$\begin{aligned}\|\boldsymbol{\pi} - \boldsymbol{\pi}^*\|_2 &\leq \frac{1}{\sqrt{\pi_{\min}^*}} \|\boldsymbol{\pi} - \boldsymbol{\pi}^*\|_{\boldsymbol{\pi}^*} \leq \frac{4\kappa^2}{\sqrt{\pi_{\min}^*}} \|\boldsymbol{\pi}^{*\top} \mathbf{E}\|_{\boldsymbol{\pi}^*} \leq 4\kappa^{2.5} \|\boldsymbol{\pi}^{*\top} \mathbf{E}\|_2 \\ &\leq 4\kappa^{2.5} \|\mathbf{E}\| \|\boldsymbol{\pi}^*\|_2.\end{aligned}$$

Here, the first inequality comes from (ii), the second inequality is a consequence of (3.76), whereas the third one results from (i). The proof is then completed by applying the high-probability bound $\|\mathbf{E}\| \lesssim \sqrt{(\log n)/n}$ derived in Lemma 3.13. \square

3.6.4 Proof of auxiliary lemmas

Before delving into the proof, we state a general comparison theorem, which is attributed to Diaconis and Saloff-Coste (1993), that relates the spectral gap of a reversible Markov chain with that of another (possibly more tractable) reversible chain. We refer the interested reader to Negahban *et al.* (2016, Lemma 6) for a proof of the following result.

Lemma 3.15. Consider two reversible Markov chains over the state space $\{1, 2, \dots, n\}$. Let \mathbf{P} and $\boldsymbol{\pi}$ (resp. $\tilde{\mathbf{P}}$ and $\tilde{\boldsymbol{\pi}}$) denote the transition matrix and the stationary distribution of the first (resp. second) chain. In addition, set

$$\alpha := \min_{i,j} \frac{\pi_i P_{i,j}}{\tilde{\pi}_i \tilde{P}_{i,j}}, \quad \text{and} \quad \beta := \max_i \frac{\pi_i}{\tilde{\pi}_i}.$$

Then one has

$$\frac{1 - \max \{\lambda_2(\mathbf{P}), -\lambda_n(\mathbf{P})\}}{1 - \max \{\lambda_2(\tilde{\mathbf{P}}), -\lambda_n(\tilde{\mathbf{P}})\}} \geq \frac{\alpha}{\beta}.$$

Armed with this comparison lemma, we are ready to present the proof of Lemma 3.12.

Proof of Lemma 3.12. In order to control the spectral gap with the aid of Lemma 3.15, we construct an auxiliary transition matrix

$$\mathbf{Q}^* = \frac{1}{n} \mathbf{1} \mathbf{1}^\top,$$

which clearly corresponds to a reversible Markov chain with stationary distribution $\mathbf{u}^* = (1/n) \cdot \mathbf{1}$. The eigengap of this newly constructed reversible Markov chain is

$$1 - \max \{\lambda_2(\mathbf{Q}^*), -\lambda_n(\mathbf{Q}^*)\} = 1,$$

since $\lambda_2(\mathbf{Q}^*) = \lambda_n(\mathbf{Q}^*) = 0$. Therefore, we only need to bound α and β .

Recalling the construction of \mathbf{P}^* in (3.73), we can straightforwardly check that

$$\pi_i^* P_{i,j}^* = \frac{1}{n} \frac{\pi_i^* \pi_j^*}{\pi_i^* + \pi_j^*} \geq \frac{1}{2n} \min\{\pi_i^*, \pi_j^*\}$$

for every $i \neq j$, and in addition,

$$\pi_i^* P_{i,i}^* = \pi_i^* \left[1 - \sum_{j:j \neq i} \frac{1}{n} \frac{w_j^*}{w_i^* + w_j^*} \right] \geq \pi_i^* \left[1 - \sum_{j:j \neq i} \frac{1}{n} \right] = \frac{1}{n} \pi_i^*.$$

Combining the previous two inequalities, we obtain

$$\min_{i,j} (\pi_i^* P_{i,j}^*) \geq \frac{1}{2n} \min_{1 \leq i \leq n} \pi_i^* = \frac{1}{2n\kappa} \max_{1 \leq i \leq n} \pi_i^* \geq \frac{1}{2n^2\kappa},$$

where the last relation holds since $\max_i \pi_i^* \geq \frac{1}{n} \sum_i \pi_i^* = \frac{1}{n}$. This together with the construction of \mathbf{Q}^* further leads to

$$\alpha := \min_{i,j} \frac{\pi_i^* P_{i,j}^*}{u_i^* Q_{i,j}^*} = n^2 \min_{i,j} (\pi_i^* P_{i,j}^*) \geq \frac{1}{2\kappa}.$$

In regard to β , it is seen that

$$\beta := \max_i \frac{\pi_i^*}{u_i^*} = n \max_i \pi_i^* = n\kappa \min_i \pi_i^* \leq \kappa,$$

where the final inequality follows since $\min_i \pi_i \leq \frac{1}{n} \sum_i \pi_i = \frac{1}{n}$. With the preceding bounds on α and β in place, Lemma 3.15 informs us that

$$\frac{1 - \max \{\lambda_2(\mathbf{P}^*), -\lambda_n(\mathbf{P}^*)\}}{1 - \max \{\lambda_2(\mathbf{Q}^*), -\lambda_n(\mathbf{Q}^*)\}} \geq \frac{\alpha}{\beta} \geq \frac{1}{2\kappa^2}.$$

This together with the aforementioned eigengap for \mathbf{Q}^* establishes the advertised result.

Proof of Lemma 3.13. Let $\mathbf{D} := \text{diag}(\sqrt{\pi_1^*}, \dots, \sqrt{\pi_n^*})$. We have seen from the proof in Section 2.5.3 (cf. (2.34)) that

$$\|\mathbf{E}\|_{\pi^*} = \|\mathbf{D}\mathbf{E}\mathbf{D}^{-1}\| \leq \|\mathbf{D}\| \|\mathbf{E}\| \|\mathbf{D}^{-1}\| = \frac{\sqrt{\max_i \pi_i^*}}{\sqrt{\min_i \pi_i^*}} \|\mathbf{E}\| = \sqrt{\kappa} \|\mathbf{E}\|,$$

where κ is defined in (3.70). Therefore, it suffices to bound $\|\mathbf{E}\|$.

By construction of \mathbf{P} and \mathbf{P}^* (see (3.72) and (3.73)), we see that

$$E_{i,j} = P_{i,j} - P_{i,j}^* = \frac{1}{n} (y_{i,j} - \mathbb{E}[y_{i,j}]) \quad (3.78)$$

for any $i \neq j$. In addition, for all $1 \leq i \leq n$, it follows that

$$E_{i,i} = P_{i,i} - P_{i,i}^* = - \sum_{j:j \neq i} E_{i,j} = - \frac{1}{n} \sum_{j:j \neq i} (y_{i,j} - \mathbb{E}[y_{i,j}]). \quad (3.79)$$

In view of these identities, we shall decompose the matrix \mathbf{E} into three parts: the upper triangular part (denoted by $\mathbf{E}_{\text{upper}}$), the diagonal part (denoted by \mathbf{E}_{diag}), and the lower triangular part (denoted by $\mathbf{E}_{\text{lower}}$). Clearly, the triangle inequality gives

$$\|\mathbf{E}\| \leq \|\mathbf{E}_{\text{upper}}\| + \|\mathbf{E}_{\text{diag}}\| + \|\mathbf{E}_{\text{lower}}\|. \quad (3.80)$$

In the sequel, we deal with these three terms separately.

Let us start with the diagonal part \mathbf{E}_{diag} . In view of the definition of the spectral norm, we know that

$$\|\mathbf{E}_{\text{diag}}\| = \max_{1 \leq i \leq n} |E_{i,i}| = \max_{1 \leq i \leq n} \left| \sum_{j:j \neq i} E_{i,j} \right|,$$

where the last relation arises from (3.79). Fix any i , then it is easily seen that $\sum_{j:j \neq i} E_{i,j}$ is a sum of independent zero-mean random variables $\{E_{i,j}\}$, which can be controlled via the Bernstein inequality. Specifically, observe that

$$\max_{j:j \neq i} |E_{i,j}| = \max_{j:j \neq i} \frac{1}{n} |y_{i,j} - \mathbb{E}[y_{i,j}]| \leq \frac{1}{n} =: B_1$$

and, in addition,

$$v_1 := \sum_{j:j \neq i} \mathbb{E}[E_{i,j}^2] = \frac{1}{n^2} \sum_{j:j \neq i} \text{Var}(y_{i,j}) \leq \frac{1}{n},$$

where the last inequality follows since the variance of a Bernoulli random variable is no larger than 1. Apply the Bernstein inequality (cf. Corollary 3.3) and the union bound over $1 \leq i \leq n$ to demonstrate that

$$\begin{aligned} \|\mathbf{E}_{\text{diag}}\| &= \max_{1 \leq i \leq n} \left| \sum_{j:j \neq i} E_{i,j} \right| \lesssim \sqrt{v_1 \log n} + B_1 \log n \\ &\lesssim \sqrt{\frac{\log n}{n}} + \frac{\log n}{n} \asymp \sqrt{\frac{\log n}{n}} \end{aligned}$$

holds with probability at least $1 - O(n^{-8})$.

We now move on to the upper triangular part $\mathbf{E}_{\text{upper}}$, whose entries $\{E_{i,j}\}_{i < j}$ are independent. Invoking Theorem 3.4 (see the remark about asymmetric version right after Theorem 3.4) with the bounds on B_1 and v_1 established above, we arrive at

$$\|\mathbf{E}_{\text{upper}}\| \lesssim \sqrt{v_1} + B_1 \log n \lesssim \sqrt{\frac{1}{n}} + \frac{\log n}{n} \asymp \sqrt{\frac{1}{n}}$$

with probability at least $1 - O(n^{-8})$. Similar arguments lead to the same upper bound on $\|\mathbf{E}_{\text{lower}}\|$, which we omit for brevity.

Substituting the upper bounds on $\|\mathbf{E}_{\text{diag}}\|$, $\|\mathbf{E}_{\text{upper}}\|$ and $\|\mathbf{E}_{\text{lower}}\|$ into (3.80), we immediately establish the desired bound.

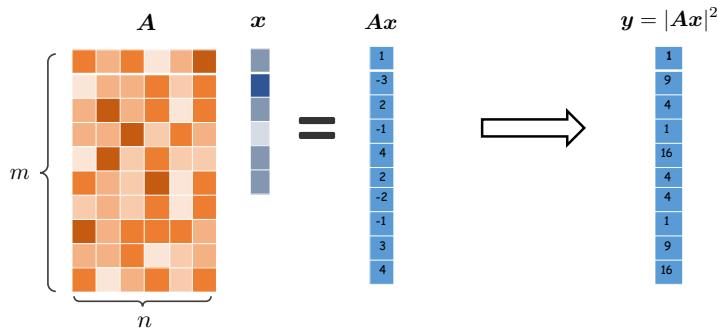


Figure 3.2: Illustration of phase retrieval and solving quadratic systems of equations, where only the intensities of linear measurements are collected. Here, $A = [a_1, \dots, a_m]^\top$, and $|z|^2 := [|z_1|^2, \dots, |z_m|^2]^\top$ for any vector $z = [z_i]_{1 \leq i \leq m}$.

3.7 Phase retrieval and solving quadratic systems of equations

Phase retrieval is a fundamental problem arising in numerous imaging applications such as X-ray crystallography, diffraction imaging, and so on (Fienup, 1982; Shechtman *et al.*, 2015; Candès *et al.*, 2013; Candès *et al.*, 2015a; Jaganathan *et al.*, 2016). In physics, phase retrieval is concerned with estimating a specimen by observing the intensities (or squared modulus) of the diffracted waves scattered by the object without knowing their phases. The advent of this problem is attributed to the physical limitation that the optical sensors are unable to record the phases of the diffracted waves. Put another way, in phase retrieval, we only have access to measurements that are quadratic functions of the object of interest, and aim at estimating the unknown object up to global phase. This gives rise to the problem of solving quadratic systems of equations, to be formulated below.

3.7.1 Problem formulation and assumptions

Suppose that we are interested in reconstructing an unknown signal $\mathbf{x}^* \in \mathbb{R}^n$, but only have access to a collection of m quadratic measurements on the linear combinations of its entries as follows:

$$y_i = (\mathbf{a}_i^\top \mathbf{x}^*)^2, \quad 1 \leq i \leq m, \quad (3.81)$$

where $\mathbf{a}_i = [a_{i,1}, \dots, a_{i,n}]^\top \in \mathbb{R}^n$ is the design vector known *a priori*. See Figure 3.2 for an illustration of this measurement model. The question is: when can we hope to reconstruct \mathbf{x}^* , in an accurate and efficient fashion, on the basis of these nonlinear equations?

As is well known, solving quadratic systems of equations is, in general, NP hard.¹ Additional assumptions are therefore needed to enable tractable recovery. Here, we adopt a Gaussian design model commonly studied in the literature.

Assumption 3.3. The design vectors $\{\mathbf{a}_i\}_{1 \leq i \leq m}$ are independently generated obeying $\mathbf{a}_i \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$.

3.7.2 Algorithm

The Gaussian design model (cf. Assumption 3.3) allows meaningful estimation of the unknown object \mathbf{x}^* via the (by now) familiar spectral method. Let us start by arranging the data into the following matrix

$$\mathbf{M} := \frac{1}{m} \sum_{i=1}^m y_i \mathbf{a}_i \mathbf{a}_i^\top = \frac{1}{m} \sum_{i=1}^m (\mathbf{a}_i^\top \mathbf{x}^*)^2 \mathbf{a}_i \mathbf{a}_i^\top, \quad (3.82)$$

which can be viewed as a weighted sample covariance matrix of the design vectors $\{\mathbf{a}_i\}$. The spectral method then estimates \mathbf{x}^* by

$$\mathbf{x} = \sqrt{\frac{\lambda_1}{3}} \mathbf{u}_1, \quad (3.83)$$

where \mathbf{u}_1 (resp. $\lambda_1 = \lambda_1(\mathbf{M})$) indicates the leading eigenvector (resp. eigenvalue) of the matrix \mathbf{M} . This simple approach has been suggested for phase retrieval since the work of Netrapalli *et al.* (2015).

To explain the rationale of this approach, it is instrumental to look at the mean of \mathbf{M} under Assumption 3.3. Specifically, simple calculation (which we include at the end of this subsection) gives

$$\mathbf{M}^* := \mathbb{E}[\mathbf{M}] = \mathbb{E}[(\mathbf{a}_i^\top \mathbf{x}^*)^2 \mathbf{a}_i \mathbf{a}_i^\top] = 2\mathbf{x}^* \mathbf{x}^{*\top} + \|\mathbf{x}^*\|_2^2 \mathbf{I}_n. \quad (3.84)$$

It is self-evident that (a) the leading eigenvector of \mathbf{M}^* is precisely given by $\pm \mathbf{x}^* / \|\mathbf{x}^*\|_2$, and (b) the leading eigenvalue of \mathbf{M}^* is given by

¹See the reduction to the NP-hard stone problem in Chen and Candès (2017).

$3\|\mathbf{x}^*\|_2^2$ by (3.84). From now on, we shall set

$$\mathbf{u}_1^* := \mathbf{x}^* / \|\mathbf{x}^*\|_2, \quad \text{and} \quad \lambda_1^* := 3\|\mathbf{x}^*\|_2^2, \quad (3.85)$$

which implies $\mathbf{x}^* = \sqrt{\lambda_1^*/3} \mathbf{u}_1^*$ and hence explains the estimator constructed in (3.83). The above argument further hints that: the spectral estimate \mathbf{x} converges to the ground truth $\pm \mathbf{x}^*$ in the large-sample limit with $m \rightarrow \infty$ (so that $\mathbf{M} \rightarrow \mathbb{E}[\mathbf{M}] = \mathbf{M}^*$). The question, however, boils down to where this algorithm stands in the more realistic finite-sample scenario.

Remark 3.9. The expression (3.85) indicates that $\mathbf{x}^* = \|\mathbf{x}^*\|_2 \mathbf{u}_1^*$. From the law of large numbers, one expects

$$\frac{1}{m} \sum_{i=1}^m y_i \rightarrow \mathbb{E}[y_i] = \mathbb{E}[(\mathbf{a}^\top \mathbf{x}^*)^2] = \|\mathbf{x}^*\|_2^2,$$

with probability approaching one. Thus, an alternative estimator is

$$\hat{\mathbf{x}} = \left(\frac{1}{m} \sum_{i=1}^m y_i \right)^{1/2} \mathbf{u}_1. \quad (3.86)$$

Derivation of (3.84). The (i, j) -th entry of $\mathbb{E}[\mathbf{M}]$ is given by

$$\mathbb{E}[M_{j,k}] = \mathbb{E}[((\mathbf{a}_i^\top \mathbf{x}^*)^2 \mathbf{x}^* \mathbf{x}^{*\top})_{j,k}] = \mathbb{E}[(a_{i,1} x_{i,1}^* + \dots + a_{i,n} x_{i,n}^*)^2 a_{i,j} a_{i,k}].$$

Expanding terms and using the moments of Gaussian variables yield

$$\begin{aligned} \mathbb{E}[M_{j,k}] &= \mathbb{E}[2a_{i,j}^2 a_{i,k}^2 x_{i,j}^* x_{i,k}^*] = 2x_{i,j}^* x_{i,k}^* \quad \text{if } j \neq k; \\ \mathbb{E}[M_{j,j}] &= \mathbb{E}[a_{i,j}^4 (x_{i,j}^*)^2] + \sum_{l: l \neq j} \mathbb{E}[a_{i,l}^2 a_{i,j}^2 (x_{i,l}^*)^2] = 3(x_{i,j}^*)^2 + \sum_{l: l \neq j} (x_{i,l}^*)^2 \\ &= 2(x_{i,j}^*)^2 + \|\mathbf{x}^*\|_2^2. \end{aligned}$$

Putting these together leads to the expression (3.84).

3.7.3 Performance guarantees

Developing theoretical support for the aforementioned spectral method hinges upon characterizing the proximity of λ_1 and λ_1^* and that of \mathbf{u}_1 and \mathbf{u}_1^* , both of which rely largely on bounding $\mathbf{M} - \mathbf{M}^*$. In what follows, we start by controlling $\|\mathbf{M} - \mathbf{M}^*\|$, with the proof postponed to Section 3.7.5.

Lemma 3.16. Consider the settings in Section 3.7.1. There exist sufficiently large constants $c, C > 0$ such that if $m \geq Cn \log^3 m$, then with probability at least $1 - O(m^{-10})$ one has

$$\|\mathbf{M} - \mathbf{M}^*\| \leq c \sqrt{\frac{n \log^3 m}{m}} \|\mathbf{x}^*\|_2^2 \leq \frac{1}{10} \|\mathbf{x}^*\|_2^2. \quad (3.87)$$

Remark 3.10. The sample size requirement can be further relaxed to $m \geq Cn \log n$ with a more careful treatment (Candès *et al.*, 2015a; Ma *et al.*, 2020). For the sake of conciseness, however, we do not strive to shave the log factors here.

With the above bound in mind, we are ready to characterize the statistical accuracy of the spectral method for phase retrieval.

Theorem 3.17. Suppose the assumptions of Lemma 3.16 hold, then with probability at least $1 - O(m^{-10})$, the following holds

$$\min\{\|\mathbf{x} - \mathbf{x}^*\|_2, \|\mathbf{x} + \mathbf{x}^*\|_2\} \leq 3c \sqrt{\frac{n \log^3 m}{m}} \|\mathbf{x}^*\|_2.$$

As can be seen from Theorem 3.17, when the number m of measurements obeys $m \gg n \log^3 m$, the relative accuracy of the spectral estimates (i.e., $\min\{\|\mathbf{x} - \mathbf{x}^*\|_2, \|\mathbf{x} + \mathbf{x}^*\|_2\}/\|\mathbf{x}^*\|_2$) becomes considerably smaller than 1, thus indicating consistent estimation. This should be contrasted with the minimax lower bounds derived in the literature (Cai and Zhang, 2015; Eldar and Mendelson, 2014), which assert that no estimator can achieve a vanishingly small relative estimation error if m is orderwise smaller than n . All this corroborates the power of spectral methods for solving the phase retrieval problem.

Proof of Theorem 3.17. Lemma 3.16 and Weyl's inequality (see Lemma 2.2) yield

$$|\lambda_1 - \lambda_1^*| \leq \|\mathbf{M} - \mathbf{M}^*\| \leq c \sqrt{\frac{n \log^3 m}{m}} \|\mathbf{x}^*\|_2^2 \leq \|\mathbf{x}^*\|_2^2, \quad (3.88)$$

As a result, by using $\lambda_1^* = 3\|\mathbf{x}^*\|_2^2$, we have

$$\lambda_1 \geq \lambda_1^* - \|\mathbf{x}^*\|_2^2 = 3\|\mathbf{x}^*\|_2^2 - \|\mathbf{x}^*\|_2^2 = 2\|\mathbf{x}^*\|_2^2. \quad (3.89)$$

In addition, note that $\lambda_1^* = \lambda_1(\mathbf{M}^*) = 3\|\mathbf{x}^*\|_2^2$ and $\lambda_i(\mathbf{M}^*) = \|\mathbf{x}^*\|_2^2$ for all $i \geq 2$. The bound (3.87) on $\mathbf{M} - \mathbf{M}^*$ indicates that

$$\|\mathbf{M} - \mathbf{M}^*\| \leq (1 - 1/\sqrt{2})[\lambda_1(\mathbf{M}^*) - \lambda_2(\mathbf{M}^*)],$$

which allows one to invoke the Davis-Kahan sin Θ theorem (cf. Corollary 2.8) to obtain

$$\text{dist}(\mathbf{u}_1, \mathbf{u}_1^*) \leq \frac{2\|\mathbf{M} - \mathbf{M}^*\|}{\lambda_1(\mathbf{M}^*) - \lambda_2(\mathbf{M}^*)} \leq 2c\sqrt{\frac{n \log^3 m}{m}}. \quad (3.90)$$

Without loss of generality, we shall assume $\|\mathbf{u}_1 - \mathbf{u}_1^*\|_2 = \text{dist}(\mathbf{u}_1, \mathbf{u}_1^*)$ in the sequel.

Now we are ready to control our target quantity $\text{dist}(\mathbf{x}, \mathbf{x}^*)$. In view of the definition (3.83) of \mathbf{x} , one has

$$\begin{aligned} \|\mathbf{x} - \mathbf{x}^*\|_2 &= \left\| \sqrt{\lambda_1/3} \mathbf{u}_1 - \|\mathbf{x}^*\|_2 \mathbf{u}_1^* \right\|_2 \\ &\leq \left\| \left(\sqrt{\lambda_1/3} - \|\mathbf{x}^*\|_2 \right) \mathbf{u}_1 \right\|_2 + \|\mathbf{x}^*\|_2 \|\mathbf{u}_1 - \mathbf{u}_1^*\|_2 \\ &\leq \left| \sqrt{\lambda_1/3} - \|\mathbf{x}^*\|_2 \right| + 2c\sqrt{\frac{n \log^3 m}{m}} \|\mathbf{x}^*\|_2. \end{aligned} \quad (3.91)$$

Here, the second line applies the triangle inequality, and the last line arises from the facts $\|\mathbf{u}_1\|_2 = 1$ and (3.90). It then boils down to controlling $|\sqrt{\lambda_1/3} - \|\mathbf{x}^*\|_2|$, for which (3.88) and (3.89) prove useful. A little algebra reveals that

$$\left| \sqrt{\lambda_1/3} - \|\mathbf{x}^*\|_2 \right| = \frac{1}{\sqrt{3}} \frac{|\lambda_1 - 3\|\mathbf{x}^*\|_2^2|}{\sqrt{\lambda_1} + \sqrt{3}\|\mathbf{x}^*\|_2} \leq c\sqrt{\frac{n \log^3 m}{m}} \|\mathbf{x}^*\|_2, \quad (3.92)$$

where the last relation relies on the bounds (3.88) and (3.89).

Taking collectively (3.91) and (3.92) concludes the proof.

3.7.4 Extensions

The spectral algorithm described in Section 3.7.2, while enjoying appealing statistical guarantees, is improvable in multiple aspects. In this subsection, we briefly discuss two central issues: sample efficiency and robustness against outliers.

Improving sample efficiency

Thus far, the spectral algorithm we have discussed requires the sample size to exceed $m \gtrsim n \log^3 m$. While this can be improved to $m \gtrsim n \log n$ via tighter analysis (Candès *et al.*, 2015a; Ma *et al.*, 2020), it remains suboptimal due to the presence of the log factor. What happens in the sample-starved regime where the sample size m is on the same order as the number n of unknowns? Is it possible to achieve the information-theoretic sampling limit for this problem? As it turns out, in order to attain the desired statistical accuracy in the sample-starved regime, we have to modify the standard recipe by applying appropriate preprocessing steps before forming the data matrix \mathbf{M} , as we shall explain momentarily.

Why is the algorithm in Section 3.7.2 suboptimal? Before introducing the improved spectral algorithm, we take a closer look at the lower bound of the approximation error $\|\mathbf{M} - \mathbf{M}^*\|$ for the sample-starved regime. Clearly,

$$\|\mathbf{M}\| \geq \frac{\mathbf{a}_j^\top \mathbf{M} \mathbf{a}_j}{\|\mathbf{a}_j\|_2^2} = \frac{1}{m} \sum_{i=1}^m y_i \frac{(\mathbf{a}_i^\top \mathbf{a}_j)^2}{\|\mathbf{a}_j\|_2^2} \geq \frac{1}{m} y_j \|\mathbf{a}_j\|_2^2$$

holds for any $1 \leq j \leq m$. Taking $j = i^* := \arg \max_i y_i$, we obtain

$$\|\mathbf{M}\| \geq \frac{(\max_i y_i) \|\mathbf{a}_{i^*}\|_2^2}{m}. \quad (3.93)$$

Under the i.i.d. Gaussian design, $\{y_i/\|\mathbf{x}^*\|_2^2\}_{1 \leq i \leq m}$ forms a collection of i.i.d. χ^2 random variables with 1 degree of freedom. Classical Gaussian concentration results (Ferguson, 1996; Vershynin, 2017) tell us that

$$\max_{1 \leq i \leq m} y_i = (2 + o(1)) \|\mathbf{x}^*\|_2^2 \log m, \quad \|\mathbf{a}_i\|_2^2 = (1 + o(1))n, \quad 1 \leq i \leq m$$

with probability approaching one as n grows, as long as $m = \text{poly}(n)$. Substitution into (3.93) implies that

$$\|\mathbf{M}\| \geq (2 + o(1)) \frac{n \log m}{m} \|\mathbf{x}^*\|_2^2 \gg \|\mathbf{x}^*\|_2^2$$

once $m \ll n \log m$, which combined with (3.84) further yields

$$\|\mathbf{M} - \mathbf{M}^*\| \geq \|\mathbf{M}\| - \|\mathbf{M}^*\| = \|\mathbf{M}\| - 3\|\mathbf{x}^*\|_2^2 \gg \|\mathbf{M}^*\|.$$

In other words, the deviation between \mathbf{M} and \mathbf{M}^* is not as well-controlled as desired in the regime with $m \ll n \log m$, and hence classical matrix perturbation theory (e.g., the Davis-Kahan theorem) does not support the use of the spectral algorithm based on \mathbf{M} in this case.

Spectral methods with data preprocessing. The above diagnosis suggests a natural remedy: since the culprit lies in the large influence $\max_i y_i$ has brought to bear on the leading eigenvector, it is advisable to downweight the effect of any excessively large y_i . This is precisely the key idea behind the *truncated spectral method* proposed by Chen and Candès (2017)—as well as other variations proposed thereafter—that provably improves the sample efficiency of spectral methods.

More specifically, instead of using the matrix \mathbf{M} constructed in (3.82), we resort to a properly preprocessed data matrix

$$\mathbf{M}_{\mathcal{T}} := \frac{1}{m} \sum_{i=1}^m \mathcal{T}(y_i) \mathbf{a}_i \mathbf{a}_i^\top \quad (3.94)$$

with \mathcal{T} some preprocessing function, and produce, by (3.86), an estimate

$$\mathbf{x}_{\mathcal{T}} = \left(\frac{1}{m} \sum_{i=1}^m y_i \right)^{1/2} \mathbf{u}_{1,\mathcal{T}}, \quad (3.95)$$

where $\mathbf{u}_{1,\mathcal{T}}$ denotes the leading eigenvector of $\mathbf{M}_{\mathcal{T}}$. A few representative examples of \mathcal{T} are in order.

- *mean-based truncation* (Chen and Candès, 2017):

$$\mathcal{T}(y) := y \mathbb{1}\{y \leq \alpha_1 \bar{y}\}, \quad \bar{y} := \frac{1}{m} \sum_{i=1}^m y_i, \quad (3.96)$$

where $\alpha_1 > 0$ is some sufficiently large constant;

- *median-based truncation* (Zhang *et al.*, 2016; Zhang *et al.*, 2018a):

$$\mathcal{T}(y) := y \mathbb{1}\{y \leq \alpha_2 y_{\text{med}}\}, \quad y_{\text{med}} := \text{median}(\{y_i\}), \quad (3.97)$$

where $\alpha_2 > 0$ is some sufficiently large constant;

- *orthogonality promotion* (Wang *et al.*, 2018a; Duchi and Ruan, 2019)

$$\mathcal{T}(y) := \mathbb{1}\{y \geq y_{(\alpha_3 m)}\}, \quad (3.98)$$

where $y_{(1)} \geq y_{(2)} \geq \dots \geq y_{(m)}$ denote the order statistics of $\{y_i\}$, and $0 < \alpha_3 < 1$ is some properly chosen constant;

- *optimal preprocessing* (Mondelli and Montanari, 2019; Luo *et al.*, 2019):

$$\mathcal{T}(y) := \frac{y / \bar{y} - 1}{y / \bar{y} + \sqrt{2m/n} - 1}, \quad \bar{y} := \frac{1}{m} \sum_{i=1}^m y_i. \quad (3.99)$$

Remark 3.11. To be more precise, $\mathcal{T}(y)$ depends not only on y but also some statistics about $\{y_i\}$ (e.g., empirical mean). Here, we suppress the dependency on such additional statistics mainly to simplify notation.

In words, the first two choices discard any measurement y_i that is too large (compared to the order of either the empirical mean or empirical median), the third one selects a subset of measurements that are most aligned with the unknown signal and scales their contributions to a measurement-invariant level, while the last one effectively behaves as a shrinkage operator once y_i rises above the empirical mean. The following theorem—which was first established in Chen and Candès (2017) for the version (3.96) and subsequently extended to other alternatives (Zhang *et al.*, 2016; Wang *et al.*, 2018a; Lu and Li, 2020; Mondelli and Montanari, 2019; Luo *et al.*, 2019)—confirms the effectiveness and importance of proper preprocessing in enabling order-optimal sample complexity. The interested reader is referred to these papers for the proofs.

Theorem 3.18. Consider the settings in Section 3.7. Fix any constant $\varepsilon > 0$, and suppose $m \geq c_0 n$ for some sufficiently large constant $c_0 > 0$ that is independent of n and m but possibly dependent on ε . Then the spectral estimate (3.95) equipped with the above choices of \mathcal{T} obeys

$$\min\{\|\mathbf{x}_\mathcal{T} - \mathbf{x}^*\|_2, \|\mathbf{x}_\mathcal{T} + \mathbf{x}^*\|_2\} \leq \varepsilon \|\mathbf{x}^*\|_2$$

with probability at least $1 - O(n^{-2})$, provided that the parameters $\alpha_1, \alpha_2, \alpha_3$ are suitably chosen in (3.96)–(3.98).

To demonstrate the practicability of preprocessing, we depict in Figure 3.3 the numerical performance of the mean-truncated spectral method (i.e., the choice (3.96)) in comparison to the vanilla version

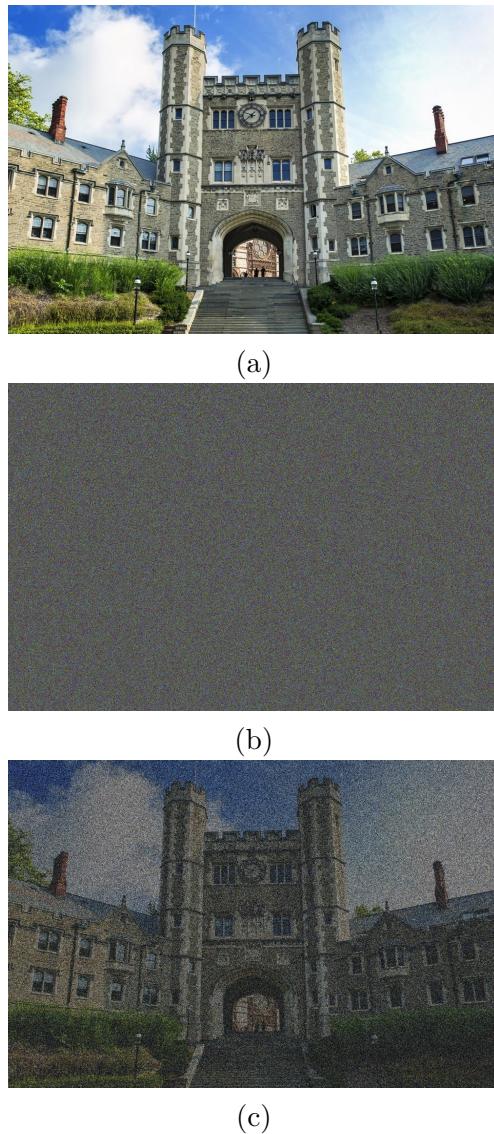


Figure 3.3: Numerical performance of spectral methods for phase retrieval under coded diffraction patterns (see Candès *et al.* (2015a) for details) when $m = 10n$. (a) the original image \mathbf{x}^* (which is 613,760-dimensional); (b) the estimate of the spectral method in Section 3.7.2; (c) the estimate of the mean-truncated spectral method (cf. (3.96)). There are in total 10 groups of measurements each of size n ; to generate each group of measurements, the entries of the signal \mathbf{x}^* are first independently multiplied by random variables uniformly over $\{1, -1, i, -i\}$, followed by an application of the discrete Fourier transform.

described in Section 3.7.2. These numerical experiments corroborate the clear advantage of proper preprocessing in the sample-starved regime.

Finally, we remark that in addition to order-wise statistical guarantees, Lu and Li (2020) further pinned down sharp characterization of the error bounds (including the pre-constants) in this sample-starved regime. Leveraging such sharp analyses, Mondelli and Montanari (2019) identified the information-theoretic optimal choice (3.99), in the sense that it leads to an estimate strictly better than a random guess (a.k.a. weak recovery) whenever it is information-theoretically possible. Luo *et al.* (2019) further showed that this choice is uniformly optimal, meaning that it leads to the smallest principal angle between \mathbf{x}_T and \mathbf{x}^* uniformly over all sampling ratios when m is on the same order of n .

Robustness vis-à-vis adversarial outliers

Another practical consideration that merits special attention is that the collected samples are sometimes susceptible to adversarial entries (due to, say, sensor failures or malicious attacks). To formulate this in more formal terms, consider the following modified measurement model (Zhang *et al.*, 2016; Hand, 2017; Hand and Voroninski, 2016):

$$y_i = \begin{cases} (\mathbf{a}_i^\top \mathbf{x}^*)^2, & i \notin \mathcal{S}_{\text{outlier}}, \\ \text{arbitrary}, & i \in \mathcal{S}_{\text{outlier}}. \end{cases} \quad (3.100)$$

Here, $\mathcal{S}_{\text{outlier}} \subseteq \{1, \dots, m\}$ represents the unknown subset of indices associated with outliers, which is of cardinality $|\mathcal{S}_{\text{outlier}}| = \alpha m$ for some $0 < \alpha < 1$. In particular, the measurements coming from $\mathcal{S}_{\text{outlier}}$ might be corrupted arbitrarily. The goal is to reliably estimate \mathbf{x}^* even when the measurements are grossly corrupted.

Unfortunately, the vanilla spectral method presented in Section 3.7.2 might not function properly even in the presence of a single outlier; for instance, if the magnitude of this outlier is excessively large, then the leading eigenvector of \mathbf{M} will be heavily biased by this outlier. As a result, the spectral method needs to be properly adjusted in order to combat the adverse effect of outliers.

To circumvent this issue, we first remind the readers of a classical finding in robust statistics (Huber, 2004): the median statistic is often

times robust against adversarial outliers. Leveraging this finding to the phase retrieval context, one might naturally employ the median of the measurements $\{y_i\}_{1 \leq i \leq m}$ as a tool to help detect any excessively large outlier. In fact, this is precisely the idea behind the median-truncated scheme presented in (3.97), whose capability in dealing with outliers has been established in Zhang *et al.* (2016) for phase retrieval and Li *et al.* (2020c) for low-rank matrix recovery.

Theorem 3.19. Consider the measurement model in (3.100), and the i.i.d. Gaussian design in Assumption 3.3. Fix any $\varepsilon > 0$. There exist some constants $c_0 > 0$ and $0 < c_1 < 1$ such that if $m \geq c_0 n$ and $\alpha \leq c_1$, then the spectral estimate (3.95) equipped with (3.97) obeys

$$\min\{\|\mathbf{x}_T - \mathbf{x}^*\|_2, \|\mathbf{x}_T + \mathbf{x}^*\|_2\} \leq \varepsilon \|\mathbf{x}^*\|_2$$

with probability at least $1 - O(n^{-2})$.

In a nutshell, Theorem 3.19 reveals that a median-truncated spectral method achieves consistent estimation even when a constant fraction of the measurements are corrupted in an arbitrary manner. All this is guaranteed to happen even when the number m of samples is on the same order as n , thus further enhancing the resilience of spectral methods in the presence of adversarial corruptions. The interested reader is referred to Zhang *et al.* (2016) for the proof of this theorem.

3.7.5 Proof of auxiliary lemmas

Proof of Lemma 3.16. Given that $\{\mathbf{a}_i\}$ is rotationally invariant, we assume without loss of generality that $\mathbf{x}^* = \mathbf{e}_1$, where \mathbf{e}_1 is the first standard basis vector. Thus, our task can be translated into bounding

$$\mathbf{M} - \mathbf{M}^* = \frac{1}{m} \sum_{i=1}^m \left\{ a_{i,1}^2 \mathbf{a}_i \mathbf{a}_i^\top - (2\mathbf{e}_1 \mathbf{e}_1^\top + \mathbf{I}_n) \right\} =: \frac{1}{m} \sum_{i=1}^m \left(\mathbf{B}_i - \mathbb{E}[\mathbf{B}_i] \right),$$

where $a_{i,1}$ denotes the first entry of the vector \mathbf{a}_i , and $\mathbf{B}_i := a_{i,1}^2 \mathbf{a}_i \mathbf{a}_i^\top$.

In order to deal with the unboundedness of Gaussian random variables, we resort to the truncated matrix Bernstein inequality, which requires us to first set a proper truncation level. In view of the Gaussianity of \mathbf{a}_i and the union bound, one has $\|\mathbf{a}_i\|_\infty \leq 5\sqrt{\log m}$ for all

$1 \leq i \leq m$ with probability at least $1 - m^{-11.5}$; on this event, one would have

$$\|\mathbf{B}_i\| = |a_{i,1}|^2 \|\mathbf{a}_i\|_2^2 \leq n \|\mathbf{a}_i\|_\infty^4 \leq 5^4 n \log^2 m. \quad (3.101)$$

Therefore, taking $L := 5^4 n \log^2 m$ leads to

$$\mathbb{P}\{\|\mathbf{B}_i\| \geq L\} \leq m^{-11.5} =: q_0.$$

Further, truncating at this level does not incur much bias; to be precise, we claim that (with the proof deferred to the end of this subsection)

$$q_1 := \|\mathbb{E}[\mathbf{B}_i \mathbb{1}\{\|\mathbf{B}_i\| < L\} - \mathbb{E}[\mathbf{B}_i]]\| \lesssim m^{-3}. \quad (3.102)$$

The next step is to characterize the variance statistic. Towards this end, it is first seen from the definition of \mathbf{B}_i that

$$\mathbb{E}[(\mathbf{B}_i - \mathbb{E}[\mathbf{B}_i])^2] \preceq \mathbb{E}[\mathbf{B}_i^2] = \mathbb{E}[a_{i,1}^4 \|\mathbf{a}_i\|_2^2 \mathbf{a}_i \mathbf{a}_i^\top]. \quad (3.103)$$

As can be easily verified, $\mathbb{E}[a_{i,1}^4 \|\mathbf{a}_i\|_2^2 \mathbf{a}_i \mathbf{a}_i^\top]$ is a diagonal matrix, whose diagonal entries obey

$$\left(\mathbb{E}[a_{i,1}^4 \|\mathbf{a}_i\|_2^2 \mathbf{a}_i \mathbf{a}_i^\top] \right)_{l,l} = \mathbb{E}[a_{i,1}^4 a_{i,l}^2 \sum_j a_{i,j}^2] = \sum_j \mathbb{E}[a_{i,1}^4 a_{i,l}^2 a_{i,j}^2] \lesssim n$$

for any $1 \leq l \leq n$, where the last relation follows from the property of standard Gaussians. This taken together with (3.103) gives

$$v := \left\| \sum_i \mathbb{E}[(\mathbf{B}_i - \mathbb{E}[\mathbf{B}_i])^2] \right\| \leq \sum_i \max_l \left| \left(\mathbb{E}[a_{i,1}^4 \|\mathbf{a}_i\|_2^2 \mathbf{a}_i \mathbf{a}_i^\top] \right)_{l,l} \right| \lesssim mn.$$

Invoking the truncated Bernstein inequality in Corollary 3.2 then yields: with probability at least $1 - O(m^{-10}) - mq_0 = 1 - O(m^{-10})$, one has

$$\begin{aligned} \|\mathbf{M} - \mathbf{M}^*\| &\lesssim \frac{1}{m} \sqrt{v \log m} + \frac{L}{m} \log m + \frac{mq_1}{m} \\ &\lesssim \sqrt{\frac{n \log m}{m}} + \frac{n \log^3 m}{m} + \frac{1}{m^3} \lesssim \sqrt{\frac{n \log^3 m}{m}} \end{aligned}$$

as desired, with the proviso that $m \gtrsim n \log^3 m$.

Proof of the inequality (3.102). We begin by employing the relation (3.101) to help modify the truncation event as follows:

$$q_1 = \|\mathbb{E}[\mathbf{B}_i \mathbb{1}\{\|\mathbf{B}_i\| < L\} - \mathbb{E}[\mathbf{B}_i]]\| = \|\mathbb{E}[\mathbf{B}_i \mathbb{1}\{\|\mathbf{B}_i\| \geq L\}]\|$$

$$\leq \|\mathbb{E}[\mathbf{B}_i \mathbb{1}\{n\|\mathbf{a}_i\|_\infty^4 \geq L\}]\| = \|\mathbb{E}[\mathbf{B}_i \mathbb{1}\{\|\mathbf{a}_i\|_\infty \geq \bar{L}\}]\|,$$

where we define $\bar{L} := (L/n)^{1/4} = 5\sqrt{\log m}$. It is easily seen that $\mathbb{E}[\mathbf{B}_i \mathbb{1}\{\|\mathbf{a}_i\|_\infty \geq \bar{L}\}]$ is a diagonal matrix and, therefore,

$$\begin{aligned} q_1 &\leq \max_l \left| \left(\mathbb{E}[\mathbf{B}_i \mathbb{1}\{\|\mathbf{a}_i\|_\infty \geq \bar{L}\}] \right)_{l,l} \right| \stackrel{(i)}{=} \max_l \mathbb{E}[a_{i,1}^2 a_{i,l}^2 \mathbb{1}\{\|\mathbf{a}_i\|_\infty \geq \bar{L}\}] \\ &\stackrel{(ii)}{\leq} 0.5 \max_l \mathbb{E}[(a_{i,1}^4 + a_{i,l}^4) \mathbb{1}\{\|\mathbf{a}_i\|_\infty \geq \bar{L}\}] = \mathbb{E}[a_{i,1}^4 \mathbb{1}\{\|\mathbf{a}_i\|_\infty \geq \bar{L}\}] \\ &\leq \mathbb{E}[a_{i,1}^4 \mathbb{1}\{|a_{i,1}| \geq \bar{L}\}] + \mathbb{E}[a_{i,1}^4 \mathbb{1}\{\max_{j \neq 1} |a_{i,j}| \geq \bar{L}\}], \end{aligned} \quad (3.104)$$

where (i) relies on the definition of \mathbf{B}_i , and (ii) comes from the AM-GM inequality. With regards to the first term of (3.104), observe that

$$\begin{aligned} \mathbb{E}[a_{i,1}^4 \mathbb{1}\{|a_{i,1}| \geq 5\sqrt{\log m}\}] &= \int_{5\sqrt{\log m}}^{\infty} \frac{2\xi^4}{\sqrt{2\pi}} e^{-\xi^2/2} d\xi \\ &\leq 2 \int_{5\sqrt{\log m}}^{\infty} \frac{e^{-\xi^2/4}}{\sqrt{2\pi}} d\xi = 2\mathbb{P}\{|a_{i,1}| \geq 2.5\sqrt{\log m}\} \lesssim \frac{1}{m^3} \end{aligned}$$

for m sufficiently large, where we have used the fact that $\xi^4 e^{-\xi^2/2} \leq e^{-\xi^2/4}$ for $\xi \geq 5\sqrt{\log m}$. Regarding the second term of (3.104), note that

$$\begin{aligned} \mathbb{E}[a_{i,1}^4 \mathbb{1}\left\{ \max_{j \neq 1} |a_{i,j}| \geq \bar{L} \right\}] &= \mathbb{E}[a_{i,1}^4] \mathbb{E}\left[\mathbb{1}\left\{ \max_{j \neq 1} |a_{i,j}| \geq \bar{L} \right\} \right] \\ &= 3\mathbb{P}\left\{ \max_{j \neq 1} |a_{i,j}| \geq 5\sqrt{\log m} \right\} \lesssim m^{-10}, \end{aligned}$$

where the first identity uses the independence between $a_{i,1}$ and $\{a_{i,j}\}_{j \neq 1}$. Substituting the preceding bounds into (3.104) establishes (3.102). \square

3.8 Matrix completion

A pressing challenge often encountered in data science applications is estimation and learning in the face of *missing data*. To elucidate how to tackle this challenge via spectral methods, we delve into the renowned matrix completion problem in this section, followed by another application called tensor completion in Section 3.9.

$$\begin{bmatrix} \checkmark & ? & ? & ? & \checkmark & ? \\ ? & ? & \checkmark & \checkmark & ? & ? \\ \checkmark & ? & ? & \checkmark & ? & ? \\ ? & ? & \checkmark & ? & ? & \checkmark \\ \checkmark & ? & ? & ? & ? & ? \\ ? & \checkmark & ? & ? & \checkmark & ? \\ ? & ? & \checkmark & \checkmark & ? & ? \end{bmatrix}$$

Figure 3.4: Illustration of matrix completion, where each “ \checkmark ” stands for an observed entry and each “ $?$ ” represents a missing entry.

Imagine that one observes a small subset of the entries in a large unknown matrix and seeks to fill in all missing entries. An archetypal example is collaborative filtering, where one aims to predict the users’ preferences on a collection of products based on partially revealed user-product ratings. See Figure 3.4 for an illustration. The problem, often referred to as matrix completion, is apparently ill-posed in general, as there are (much) fewer measurements than the unknowns.

Fortunately, if the matrix of interest exhibits certain low-dimensional structure, then reliable recovery becomes feasible. A commonly encountered example of this kind concerns the case when the target matrix enjoys a low-rank structure. Again, take collaborative filtering for example: the user-product rating matrix might be well explained by a relatively small number of latent factors connecting users’ preferences with products’ attributes, thus resulting in an approximately low-rank matrix. Motivated by its fundamental importance, recent years have witnessed a flurry of research activity in studying low-rank matrix completion (Candès and Recht, 2009; Keshavan *et al.*, 2010; Gross, 2011); see Chen and Chi (2018) and Davenport and Romberg (2016) for overviews of recent developments. In the sequel, we present a simple yet effective approach enabled by the spectral method, originally proposed in Achlioptas and McSherry (2007) and Keshavan *et al.* (2010).

3.8.1 Problem formulation and assumptions

Suppose that we are interested in estimating an $n_1 \times n_2$ rank- r matrix $\mathbf{M}^* = [M_{i,j}^*]_{1 \leq i \leq n_1, 1 \leq j \leq n_2}$. Without loss of generality, we assume

$$n_1 \leq n_2.$$

Denote by $\mathbf{M}^* = \mathbf{U}^* \boldsymbol{\Sigma}^* \mathbf{V}^{*\top}$ the SVD of \mathbf{M}^* , where the columns of $\mathbf{U}^* \in \mathbb{R}^{n_1 \times r}$ (resp. $\mathbf{V}^* \in \mathbb{R}^{n_2 \times r}$) are the left (resp. right) singular vectors of \mathbf{M}^* , and $\boldsymbol{\Sigma}^*$ is a diagonal matrix whose diagonal entries are the singular values of \mathbf{M}^* . We define the condition number of the matrix \mathbf{M}^* to be $\kappa := \sigma_1(\mathbf{M}^*)/\sigma_r(\mathbf{M}^*)$.

To capture the presence of missing data, we introduce an index subset $\Omega \subseteq [n_1] \times [n_2]$, such that each entry $M_{i,j}^*$ is observed if and only if $(i, j) \in \Omega$. The goal is to reconstruct the singular subspaces \mathbf{U}^* and \mathbf{V}^* , as well as the full matrix \mathbf{M}^* , based on entries observed over the sampling set Ω .

Random sampling. Apparently, not all sampling patterns admit reliable estimation. For instance, if Ω contains only entries in the top half of the matrix, then there is in general no hope to predict the bottom half of the matrix. In order to allow for meaningful matrix completion, this monograph focuses on a natural random observation model commonly adopted in the literature, as formulated below.

Assumption 3.4 (Random sampling). Each entry of \mathbf{M}^* is observed independently with probability $0 < p < 1$, namely, each $(i, j) \in [n_1] \times [n_2]$ is included in Ω independently with probability p .

Under this model, we shall view the expected number of observed entries—namely, $p n_1 n_2$ —as the sample size. In truth, as long as p is not overly small, the number of observed entries is expected to concentrate around its mean $p n_1 n_2$.

Incoherence conditions. Caution needs to be exercised, however, that the random sampling model alone does not guarantee effective recovery of an arbitrary low-rank matrix \mathbf{M}^* . Consider, for example, the following rank-1 matrix \mathbf{M}^* containing a single nonzero entry:

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

If $p = o(1)$, then with probability $1 - p = 1 - o(1)$, the sampling pattern will fail to include the nonzero entry $M_{1,1}^*$, thus ruling out the possibility of faithful matrix recovery. Consequently, one needs to make sure that the sampling pattern does not suppress too much useful information. Towards this end, the pioneering work Candès and Recht (2009) and Candès and Tao (2010) singled out an incoherence parameter that plays a vital role.

Definition 3.1. The incoherence parameter μ of the matrix $\mathbf{M}^* \in \mathbb{R}^{n_1 \times n_2}$ is defined as

$$\mu := \max \left\{ \frac{n_1 \|\mathbf{U}^*\|_{2,\infty}^2}{r}, \frac{n_2 \|\mathbf{V}^*\|_{2,\infty}^2}{r} \right\}.$$

Remark 3.12. Recognizing the following basic relation

$$\frac{r}{n_1} = \frac{1}{n_1} \|\mathbf{U}^*\|_{\text{F}}^2 \leq \|\mathbf{U}^*\|_{2,\infty}^2 \leq \|\mathbf{U}^*\|^2 = 1$$

and an analogous one for \mathbf{V}^* , we have $1 \leq \mu \leq \max\{n_1, n_2\}/r = n_2/r$.

In words, a small μ indicates that the energy of the singular vectors is spread out across different elements, namely, the singular subspace of \mathbf{M}^* is not too “aligned” with any of the standard basis vectors, thus ensuring that entrywise observations provide somewhat equalized information about the full spectrum of \mathbf{M}^* . The following lemma summarizes a few immediate consequences of this definition, with the proof deferred to Section 3.8.4.

Lemma 3.20. Assume that $\mathbf{M}^* \in \mathbb{R}^{n_1 \times n_2}$ is μ -incoherent. Then the following relations hold

$$\|\mathbf{M}^*\|_{2,\infty} \leq \sqrt{\mu r/n_1} \|\mathbf{M}^*\|; \quad \|\mathbf{M}^{*\top}\|_{2,\infty} \leq \sqrt{\mu r/n_2} \|\mathbf{M}^*\|; \quad (3.105a)$$

$$\|\mathbf{M}^*\|_{\infty} \leq \mu r \|\mathbf{M}^*\| / \sqrt{n_1 n_2}. \quad (3.105b)$$

Additional notation. We find it convenient to introduce a Euclidean projection operator $\mathcal{P}_\Omega : \mathbb{R}^{n_1 \times n_2} \mapsto \mathbb{R}^{n_1 \times n_2}$ such that

$$[\mathcal{P}_\Omega(\mathbf{A})]_{i,j} = \begin{cases} A_{i,j}, & \text{if } (i,j) \in \Omega \\ 0, & \text{else} \end{cases} \quad (3.106)$$

for any matrix $\mathbf{A} = [A_{i,j}] \in \mathbb{R}^{n_1 \times n_2}$. With this notation in place, matrix completion amounts to recovering \mathbf{M}^* on the basis of $\mathcal{P}_\Omega(\mathbf{M}^*)$.

3.8.2 Algorithm

To apply the spectral method, the first step is to form a reasonable approximation \mathbf{M} of the unknown matrix \mathbf{M}^* . By virtue of the random sampling model (cf. Assumption 3.4), a candidate approximation can be obtained from the observed data matrix via inverse probability weighting:

$$\mathbf{M} := p^{-1} \mathcal{P}_\Omega(\mathbf{M}^*). \quad (3.107)$$

The rationale is that \mathbf{M} forms an unbiased estimate of the ground truth, namely,

$$\mathbb{E}[\mathbf{M}] = \mathbf{M}^*,$$

where the expectation is taken over the randomness in Ω .

As a result, the proposed spectral method proceeds by computing the rank- r SVD $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$ of the matrix \mathbf{M} constructed in (3.107), and employing $\mathbf{U} \in \mathbb{R}^{n_1 \times r}$, $\mathbf{V} \in \mathbb{R}^{n_2 \times r}$ and $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$ as estimates of \mathbf{U}^* , \mathbf{V}^* and \mathbf{M}^* , respectively.

3.8.3 Performance guarantees

As before, whether the subspace \mathbf{U} (resp. \mathbf{V}) is close to \mathbf{U}^* (resp. \mathbf{V}^*) relies crucially on the size of the perturbation $\|\mathbf{M} - \mathbf{M}^*\|$. Therefore, we begin by developing an upper bound on this quantity; the proof is based on the matrix Bernstein inequality and is postponed to Section 3.8.4.

Lemma 3.21. Consider the settings in Section 3.8.1. Suppose that $n_2 p \geq C \mu r \log n_2$ for some constant $C > 0$. Then with probability at least $1 - O(n_2^{-10})$, the matrix \mathbf{M} constructed in (3.107) obeys

$$\|\mathbf{M} - \mathbf{M}^*\| \lesssim \sqrt{\frac{\mu r \log n_2}{n_1 p}} \|\mathbf{M}^*\|.$$

With this perturbation bound in place, we are equipped to apply Wedin's $\sin \Theta$ theorem to obtain the following results. The condition on the sample size in Theorem 3.22 is stronger than that in Lemma 3.21, as we need to control the eigengap in the following theorem.

Theorem 3.22. Consider the settings in Section 3.8.1. Suppose that $n_1 p \geq C_1 \kappa^2 \mu r \log n_2$ for some sufficiently large constant $C_1 > 0$. Then with probability exceeding $1 - O(n_2^{-10})$,

$$\max \left\{ \text{dist}(\mathbf{U}, \mathbf{U}^*), \text{dist}(\mathbf{V}, \mathbf{V}^*) \right\} \lesssim \kappa \sqrt{\frac{\mu r \log n_2}{n_1 p}}.$$

Proof. As a direct consequence of Lemma 3.21, one has

$$\|\mathbf{M} - \mathbf{M}^*\| \lesssim \sqrt{\frac{\mu r \log n_2}{n_1 p}} \|\mathbf{M}^*\| \leq \left(1 - \frac{1}{\sqrt{2}}\right) \sigma_r(\mathbf{M}^*),$$

provided that $n_1 p \geq C_1 \kappa^2 \mu r \log n_2$ for some large enough constant $C_1 > 0$. Apply Wedin's theorem (cf. (2.26)) and Lemma 3.21 to obtain

$$\max \left\{ \text{dist}(\mathbf{U}, \mathbf{U}^*), \text{dist}(\mathbf{V}, \mathbf{V}^*) \right\} \leq \frac{2\|\mathbf{M} - \mathbf{M}^*\|}{\sigma_r(\mathbf{M}^*)} \lesssim \kappa \sqrt{\frac{\mu r \log n_2}{n_1 p}}$$

as claimed. \square

As an important implication of Theorem 3.22, once the sample size exceeds

$$pn_1 n_2 \gg \kappa^2 \mu r n_2 \log n_2,$$

then the spectral estimate achieves consistent estimation in the sense that

$$\max \left\{ \text{dist}(\mathbf{U}, \mathbf{U}^*), \text{dist}(\mathbf{V}, \mathbf{V}^*) \right\} = o(1).$$

Given that $pn_1 n_2 \gtrsim \mu n_2 r \log n_2$ is an information-theoretic sampling requirement for reliable matrix completion when $r = o(n_1 / \log n_2)$ (Candès and Tao, 2010), Theorem 3.22 confirms the near optimality of spectral methods—in terms of the scaling with n_1 , n_2 and p —when it comes to consistent subspace estimation.

Before moving forward to the proof, we further characterize the statistical accuracy of $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$ in estimating the unknown matrix \mathbf{M}^* .

Accomplishing this only requires Lemma 3.21, without any need of the singular subspace perturbation theory. This result will also come in handy when we turn to discussing entrywise estimation accuracy in Chapter 4.

Theorem 3.23. Consider the settings in Section 3.8.1. Suppose that $n_2 p \geq C \mu r \log n_2$ for some sufficiently large constant $C > 0$. Then with probability at least $1 - O(n_2^{-10})$, one has

$$\|\mathbf{U}\Sigma\mathbf{V}^\top - \mathbf{M}^*\|_{\text{F}} \lesssim \sqrt{\frac{\mu r^2 \log n_2}{n_1 p}} \|\mathbf{M}^*\|.$$

Proof. First, note

$$\|\mathbf{U}\Sigma\mathbf{V}^\top - \mathbf{M}^*\| \leq \|\mathbf{U}\Sigma\mathbf{V}^\top - \mathbf{M}\| + \|\mathbf{M} - \mathbf{M}^*\| \leq 2\|\mathbf{M} - \mathbf{M}^*\|,$$

where the first inequality comes from the triangle inequality, and the second inequality follows from the fact that $\mathbf{U}\Sigma\mathbf{V}^\top$ is the best rank- r approximation to \mathbf{M} , i.e.,

$$\|\mathbf{U}\Sigma\mathbf{V}^\top - \mathbf{M}\| = \min_{\mathbf{Z}: \text{rank}(\mathbf{Z}) \leq r} \|\mathbf{Z} - \mathbf{M}\| \leq \|\mathbf{M} - \mathbf{M}^*\|.$$

Additionally, it is observed that $\mathbf{U}\Sigma\mathbf{V}^\top - \mathbf{M}^*$ has rank at most $2r$, which implies

$$\|\mathbf{U}\Sigma\mathbf{V}^\top - \mathbf{M}^*\|_{\text{F}} \leq \sqrt{2r} \|\mathbf{U}\Sigma\mathbf{V}^\top - \mathbf{M}^*\| \leq 2\sqrt{2r} \|\mathbf{M} - \mathbf{M}^*\|.$$

This combined with Lemma 3.21 immediately concludes the proof. \square

3.8.4 Proof of auxiliary lemmas

Proof of Lemma 3.20. First of all, the $\|\cdot\|_{2,\infty}$ norm of \mathbf{M}^* can be upper bounded by

$$\|\mathbf{M}^*\|_{2,\infty} = \|\mathbf{U}^* \Sigma^* \mathbf{V}^{*\top}\|_{2,\infty} \leq \|\mathbf{U}^*\|_{2,\infty} \|\Sigma^*\| \|\mathbf{V}^*\| \leq \sqrt{\frac{\mu r}{n_1}} \|\mathbf{M}^*\|.$$

Here, the first inequality arises from the elementary bounds $\|\mathbf{AB}\|_{2,\infty} \leq \|\mathbf{A}\|_{2,\infty} \|\mathbf{B}\|$ and $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \|\mathbf{B}\|$, whereas the last relation uses Definition 3.1, the orthonormality of \mathbf{V}^* , and identifies $\|\Sigma^*\|$ with

$\|\mathbf{M}^*\|$. The bound on $\|\mathbf{M}^{*\top}\|_{2,\infty}$ can be derived analogously and is omitted for brevity.

In addition, the matrix \mathbf{M}^* is elementwise bounded by

$$\|\mathbf{M}^*\|_\infty = \|\mathbf{U}^* \Sigma^* \mathbf{V}^{*\top}\|_\infty \leq \|\mathbf{U}^*\|_{2,\infty} \|\mathbf{V}^*\|_{2,\infty} \|\Sigma^*\| \leq \frac{\mu r}{\sqrt{n_1 n_2}} \|\mathbf{M}^*\|.$$

Here, the first inequality follows from the fact $\|\mathbf{AB}^\top\|_\infty \leq \|\mathbf{A}\|_{2,\infty} \|\mathbf{B}\|_{2,\infty}$ and the aforementioned one $\|\mathbf{AB}\|_{2,\infty} \leq \|\mathbf{A}\|_{2,\infty} \|\mathbf{B}\|$, while the last inequality again relies on Definition 3.1.

Proof of Lemma 3.21. Note that the matrix $\mathbf{E} := p^{-1} \mathcal{P}_\Omega(\mathbf{M}^*) - \mathbf{M}^*$ can be expressed as the sum of $n_1 n_2$ i.i.d. random matrices

$$\frac{1}{p} \mathcal{P}_\Omega(\mathbf{M}^*) - \mathbf{M}^* = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \underbrace{(p^{-1} \delta_{i,j} - 1) M_{i,j}^* \mathbf{e}_i \mathbf{e}_j^\top}_{=: \mathbf{X}_{i,j}}.$$

Here, $\delta_{i,j}$ (which indicates whether the (i, j) -th entry is observed) follows an independent Bernoulli distribution with parameter p , and \mathbf{e}_i stands for the i -th standard basis vector of appropriate dimensions. It is easily seen that for each (i, j) ,

$$\mathbb{E}[\mathbf{X}_{i,j}] = \mathbf{0} \quad \text{and} \quad \|\mathbf{X}_{i,j}\| \leq \frac{1}{p} \|\mathbf{M}^*\|_\infty \leq \frac{\mu r}{p \sqrt{n_1 n_2}} \|\mathbf{M}^*\|,$$

where the last relation results from the entrywise upper bound (3.105b) on \mathbf{M}^* . In order to apply the matrix Bernstein inequality (cf. Corollary 3.3), we need to control the variance statistic

$$v := \max \left\{ \left\| \sum_{i,j} \mathbb{E} [\mathbf{X}_{i,j} \mathbf{X}_{i,j}^\top] \right\|, \left\| \sum_{i,j} \mathbb{E} [\mathbf{X}_{i,j}^\top \mathbf{X}_{i,j}] \right\| \right\}.$$

Regarding the first variance term, we have

$$\begin{aligned} \sum_{i,j} \mathbb{E} [\mathbf{X}_{i,j} \mathbf{X}_{i,j}^\top] &= \sum_{i,j} \mathbb{E} [(p^{-1} \delta_{i,j} - 1)^2 (M_{i,j}^*)^2 \mathbf{e}_i \mathbf{e}_j^\top \mathbf{e}_j \mathbf{e}_i^\top] \\ &= \frac{1-p}{p} \sum_{i,j} (M_{i,j}^*)^2 \mathbf{e}_i \mathbf{e}_i^\top = \frac{1-p}{p} \sum_{i=1}^{n_1} \|M_{i,\cdot}^*\|_2^2 \mathbf{e}_i \mathbf{e}_i^\top \\ &\preceq \frac{1-p}{p} \|\mathbf{M}^*\|_{2,\infty}^2 \mathbf{I}_{n_1} \preceq \frac{\mu r}{n_1 p} \|\mathbf{M}^*\|^2 \mathbf{I}_{n_1}. \end{aligned}$$

Here, the first identity arises from the definition of $\mathbf{X}_{i,j}$, the second one calculates the variance of Bernoulli random variables, and the last line relies on the upper bound (3.105a). Similarly, the second term in the variance statistic enjoys the following characterization:

$$\sum_{i,j} \mathbb{E} [\mathbf{X}_{i,j}^\top \mathbf{X}_{i,j}] \preceq \frac{\mu r}{n_2 p} \|\mathbf{M}^*\|^2 \mathbf{I}_{n_2}.$$

Taking the above relations together and recalling that $n_1 \leq n_2$ give

$$v \leq \frac{\mu r}{n_1 p} \|\mathbf{M}^*\|^2.$$

With the above bounds in place, invoking matrix Bernstein (see Corollary 3.3) reveals that: with probability at least $1 - O(n_2^{-10})$,

$$\|\mathbf{E}\| \lesssim \sqrt{\frac{\mu r \|\mathbf{M}^*\|^2 \log n_2}{n_1 p}} + \frac{\mu r \|\mathbf{M}^*\| \log n_2}{p \sqrt{n_1 n_2}} \asymp \sqrt{\frac{\mu r \|\mathbf{M}^*\|^2 \log n_2}{n_1 p}},$$

where the last inequality is valid as long as $n_2 p \gtrsim \mu r \log n_2$.

3.9 Tensor completion

Tensor data, which can be viewed as a higher-order generalization of matrix data, are routinely used in science and engineering applications to capture multi-way interactions across variables of interest (Kolda and Bader, 2009; Sidiropoulos *et al.*, 2017; Anandkumar *et al.*, 2014). Akin to matrix completion, the problem of tensor completion aims to reconstruct a (structured) tensor when the vast majority of its entries are unobserved, a task that spans a wide spectrum of applications including visual data inpainting, harmonic retrieval, seismic data analysis, and so on (Liu *et al.*, 2012; Chen and Chi, 2014; Kreimer *et al.*, 2013); see Fig. 3.5 for an illustration.

Apparently, this task cannot possibly be accomplished without exploiting further structural assumptions on the tensor under consideration. Inspired by the success of low-rank matrix completion, we explore the case where the unknown tensor enjoys certain low-rank structure (more specifically, low canonical-polyadic (CP) rank (Kolda and Bader, 2009)). For simplicity, we concentrate on order-three tensors (namely, $\mathbf{T} = [T_{i,j,k}] \in \mathbb{R}^{n_1 \times n_2 \times n_3}$), which already capture several fundamental challenges intrinsic to tensor estimation. In addition, we take the dimensionality $n_1 = n_2 = n_3 = n$ for simplicity of presentation.

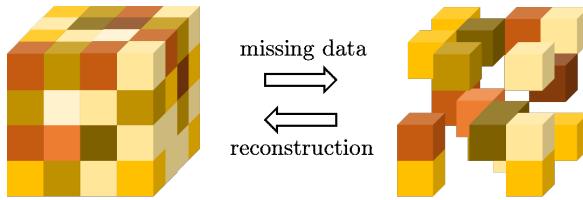


Figure 3.5: Illustration of tensor completion, where we observe partial entries of an order-three tensor.

3.9.1 Problem formulation and assumptions

Notation. Before describing our models, we introduce several notation that will be useful throughout. For any vectors $\mathbf{a} = [a_i]_{1 \leq i \leq n}$, $\mathbf{b} = [b_i]_{1 \leq i \leq n}$, $\mathbf{c} = [c_i]_{1 \leq i \leq n} \in \mathbb{R}^n$, the tensor $\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c}$ stands for an $n \times n \times n$ array whose (i, j, k) -th entry is given by $a_i b_j c_k$. Additionally, we denote by $\mathbf{a} \otimes \mathbf{b} := \begin{bmatrix} a_1 \mathbf{b} \\ \vdots \\ a_n \mathbf{b} \end{bmatrix}$ the Kronecker product between \mathbf{a} and \mathbf{b} . For any tensor $\mathbf{T} = [T_{i,j,k}]_{1 \leq i,j,k \leq n} \in \mathbb{R}^{n \times n \times n}$, we say that $\mathbf{A} = [A_{i,j}] \in \mathbb{R}^{n \times n^2}$ is the mode-1 matricization of \mathbf{T} , denoted by

$$\mathbf{A} = \text{unfold}(\mathbf{T}),$$

if $A_{i,(j-1)n+k} = T_{i,j,k}$ for all $(i, j, k) \in [n] \times [n] \times [n]$.

Models and assumptions. Suppose the unknown order-three symmetric tensor $\mathbf{T}^* = [T_{i,j,k}^*]_{1 \leq i,j,k \leq n}$ is a superposition of r ($r < n$) rank-one symmetric tensors:

$$\mathbf{T}^* = \sum_{i=1}^r \mathbf{w}_i^* \otimes \mathbf{w}_i^* \otimes \mathbf{w}_i^* \in \mathbb{R}^{n \times n \times n}, \quad (3.108)$$

where $\{\mathbf{w}_i^* \in \mathbb{R}^n\}$ represents a set of latent tensor factors. What we have available are incomplete observations of the entries of \mathbf{T}^* . The observed data can be succinctly encoded by an index subset $\Omega \subseteq [n] \times [n] \times [n]$ (called a sampling set) and a tensor $\mathbf{T} = [T_{i,j,k}]_{1 \leq i,j,k \leq n}$ as follows

$$T_{i,j,k} = \begin{cases} T_{i,j,k}^*, & \text{if } (i, j, k) \in \Omega, \\ 0, & \text{else.} \end{cases} \quad (3.109)$$

This subsection aims for an intermediate goal, namely, estimating the subspace spanned by $\{\mathbf{w}_i^*\}_{1 \leq i \leq r}$, which often serves as a crucial initial stage towards reliable completion of the whole tensor. The interested reader is referred to Montanari and Sun (2018) and Cai *et al.* (2021b) for subsequent stages of tensor completion algorithms.

Similar to the matrix completion counterpart, we explore a *random sampling* pattern such that for all $(i, j, k) \in [n] \times [n] \times [n]$,

$$(i, j, k) \in \Omega \quad \text{independently with probability } p. \quad (3.110)$$

In addition, we define for notational convenience that

$$\nu_i := \|\mathbf{w}_i^*\|_2^3, \quad \nu_{\min} := \min_{1 \leq i \leq r} \nu_i, \quad \nu_{\max} := \max_{1 \leq i \leq r} \nu_i, \quad (3.111)$$

where ν_i reflects the size of the rank-1 component $\mathbf{w}_i^* \otimes \mathbf{w}_i^* \otimes \mathbf{w}_i^*$. The condition number of \mathbf{T}^* is then defined as $\kappa := \nu_{\max}/\nu_{\min}$.

We shall also introduce several incoherence parameters as follows.

Definition 3.2. Define the incoherence parameters of \mathbf{T}^* (cf. (3.108)) as

$$\mu_1 := \max_{1 \leq i \leq r} \frac{n\|\mathbf{w}_i^*\|_{\infty}^2}{\|\mathbf{w}_i^*\|_2^2}, \quad \text{and} \quad \mu_2 := \max_{i \neq j} \frac{n|\langle \mathbf{w}_i^*, \mathbf{w}_j^* \rangle|^2}{\|\mathbf{w}_i^*\|_2^2 \|\mathbf{w}_j^*\|_2^2}. \quad (3.112)$$

Let us explain these parameters in words: small μ_1 and μ_2 reflect that (i) the energy of each tensor factor \mathbf{w}_i^* is spread out across different entries, and (ii) the factors $\{\mathbf{w}_i^*\}$ are not too correlated with each other. To simplify presentation, we set

$$\mu := \max\{\mu_1, \mu_2\}.$$

3.9.2 Algorithm

Unfortunately, it is notoriously difficult to exploit the low-rank structure—and many other low-complexity structure—efficiently in the original tensor space (Hillar and Lim, 2013). To circumvent this issue, a natural strategy thus attempts to matricize the tensor data, followed by an application of suitable low-rank matrix estimation algorithms. Specifically, let us unfold the tensor \mathbf{T}^* into an $n \times n^2$ matrix \mathbf{A}^* as follows

$$\mathbf{A}^* := \text{unfold}(\mathbf{T}^*) = \sum_{i=1}^r \mathbf{w}_i^* (\mathbf{w}_i^* \otimes \mathbf{w}_i^*)^{\top} \in \mathbb{R}^{n \times n^2}. \quad (3.113)$$

The resulting matrix \mathbf{A}^* inherits the low-rank structure, as it clearly has rank at most r . We shall also matricize the observed data as

$$\mathbf{A} := \text{unfold}(\mathbf{T}). \quad (3.114)$$

In order to estimate the subspace \mathbf{U}^* spanned by $\{\mathbf{w}_i^*\}_{1 \leq i \leq r}$ (which is the column space of \mathbf{A}^* as well), the spectral method studied here resorts to the rescaled Gram matrix $p^{-2}\mathbf{A}\mathbf{A}^\top$. As a sanity check, if there is absolutely no missing data (i.e., $p = 1$), then $p^{-2}\mathbf{A}\mathbf{A}^\top$ reduces to $\mathbf{A}^*\mathbf{A}^{*\top}$, whose column space coincides with that of \mathbf{A}^* . Turning to the scenario with missing data, a close inspection reveals that

$$\frac{1}{p^2}\mathbb{E}[\mathbf{A}\mathbf{A}^\top] = \mathbf{A}^*\mathbf{A}^{*\top} + \left(\frac{1}{p} - 1\right)\mathcal{P}_{\text{diag}}(\mathbf{A}^*\mathbf{A}^{*\top}), \quad (3.115)$$

where $\mathcal{P}_{\text{diag}}(\cdot)$ denotes the Euclidean projection onto the set of matrices with zero off-diagonal entries. This, however, makes apparent a severe issue: in the highly subsampled regime (i.e., where p is small), the diagonal components might be excessively large and non-identical, thus destroying the low-rank structure in (3.115).

To mitigate their undesirable effects, it is advisable to properly adjust the sizes of the diagonal entries (Montanari and Sun, 2018; Cai *et al.*, 2021a). As it turns out, a simple yet plausible scheme is *diagonal deletion*, which exploits only the off-diagonal part as follows

$$\mathbf{M} := \frac{1}{p^2}\mathcal{P}_{\text{off-diag}}(\mathbf{A}\mathbf{A}^\top). \quad (3.116a)$$

Here, $\mathcal{P}_{\text{off-diag}}(\cdot)$ stands for the operator that zeros out all diagonal entries of a matrix. One can easily verify that, in expectation,

$$\mathbb{E}[\mathbf{M}] = \underbrace{\mathbf{A}^*\mathbf{A}^{*\top}}_{=: \mathbf{M}^*} - \mathcal{P}_{\text{diag}}(\mathbf{A}^*\mathbf{A}^{*\top}), \quad (3.116b)$$

which stays quite close to the low-rank matrix \mathbf{M}^* as long as the diagonal entries of \mathbf{M}^* are small enough. The spectral method then proceeds by calculating the top- r eigendecomposition $\mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top$ of \mathbf{M} and returning \mathbf{U} as the subspace estimate. Here, the columns of $\mathbf{U} \in \mathbb{R}^{n \times r}$ are formed by the r leading eigenvectors of \mathbf{M} , while $\mathbf{\Lambda} \in \mathbb{R}^{r \times r}$ is a diagonal matrix containing the r leading eigenvalues.

Remark 3.13. The diagonal deletion idea has been recommended not just for tensor completion, but also for problems including but not limited to bi-clustering (Florescu and Perkins, 2016), PCA with missing data and/or heteroskedastic noise (Cai *et al.*, 2021a; Abbe *et al.*, 2020a), and contextual community detection (Abbe *et al.*, 2020a). Instead of diagonal deletion, one might also consider properly rescaling the diagonal entries based on the sampling mechanism; see, e.g., Montanari and Sun (2018), Lounici (2014), Loh and Wainwright (2012), Zhang *et al.* (2021), and Zhu *et al.* (2019).

3.9.3 Performance guarantees

The aforementioned spectral method can be analyzed by means of the ℓ_2 perturbation theory as well. As usual, this requires first controlling the size of $\mathbf{E} := \mathbf{M} - \mathbf{M}^*$, where \mathbf{M} and \mathbf{M}^* are defined in (3.116a) and (3.116b), respectively.

Lemma 3.24. Consider the settings in Section 3.9.1. There exists some universal constant $C > 0$ such that with probability at least $1 - O(n^{-7})$,

$$\|\mathbf{E}\| \leq C \left(\frac{\mu^{3/2} r \sqrt{\log n}}{n^{3/2} p} + \sqrt{\frac{\mu^2 r \log n}{n^2 p}} + \frac{\mu r}{n} \right) \nu_{\max}^2, \quad (3.117)$$

provided that $p \gtrsim \frac{\mu^{3/2} r \log^{2.5} n}{n^{3/2}}$ and that $\mu \max\{\log n, r^2 \kappa^4\} \leq c_3 n$ for some sufficiently small constant $c_3 > 0$.

In order to apply the Davis-Kahan $\sin \Theta$ theorem (cf. Corollary 2.8), another step boils down to characterizing the eigengap of the matrix $\mathbf{M}^* = \mathbf{A}^* \mathbf{A}^{*\top}$ of interest. Our result is this:

Lemma 3.25. Suppose that $\mu r^2 \kappa^4 \leq c_3 n$ for some sufficiently small constant $c_3 > 0$. Then the i -th largest eigenvalue of $\mathbf{A}^* \mathbf{A}^{*\top}$ obeys

$$\begin{aligned} \lambda_i(\mathbf{A}^* \mathbf{A}^{*\top}) &\in [\nu_{\min}^2/2, 2\nu_{\max}^2], & \text{if } 1 \leq i \leq r; \\ \lambda_i(\mathbf{A}^* \mathbf{A}^{*\top}) &= 0, & \text{if } i \geq r + 1. \end{aligned}$$

The preceding two lemmas, which will be established in Section 3.9.4, readily lead to the following statistical guarantees for the spectral method presented in Section 3.9.2.

Theorem 3.26. Consider the settings in Section 3.9.1. Suppose that

$$\mu r^2 \kappa^4 \log n \leq c_4 n \quad \text{and} \quad p \geq c_5 \frac{\mu^{3/2} \kappa^2 r \log^{2.5} n}{n^{3/2}} \quad (3.118)$$

hold for some small (resp. large) enough constant $c_4 > 0$ (resp. $c_5 > 0$). Then with probability at least $1 - O(n^{-7})$, one has

$$\text{dist}(\mathbf{U}, \mathbf{U}^*) \lesssim \frac{\mu^{3/2} \kappa^2 r \sqrt{\log n}}{n^{3/2} p} + \sqrt{\frac{\mu^2 \kappa^4 r \log n}{n^2 p}} + \frac{\mu \kappa^2 r}{n}.$$

Proof. In view of Lemmas 3.24–3.25, one would have $\|\mathbf{E}\| \leq (1 - 1/\sqrt{2})\lambda_r(\mathbf{M}^*)$ under Condition (3.118). Corollary 2.8 combined with Lemma 3.24 then tells us that, with probability at least $1 - O(n^{-7})$,

$$\text{dist}(\mathbf{U}, \mathbf{U}^*) \leq \frac{2\|\mathbf{E}\|}{\lambda_r(\mathbf{M}^*)} \lesssim \frac{\left(\frac{\mu^{3/2} r \sqrt{\log n}}{n^{3/2} p} + \sqrt{\frac{\mu^2 r \log n}{n^2 p}} + \frac{\mu r}{n} \right) \nu_{\max}^2}{\nu_{\min}^2}$$

as desired. \square

Theorem 3.26 is noteworthy for its implication on the sample complexity. To be precise, consider, for simplicity, the scenario where $r, \mu, \kappa = O(1)$. In order to achieve consistent estimation in the sense that $\text{dist}(\mathbf{U}, \mathbf{U}^*) = o(1)$, it suffices for the sample size—which sharply concentrates around $n^3 p$ under our model—to exceed

$$n^3 p \gtrsim n^{3/2} \text{poly log}(n).$$

The careful reader might immediately remark that this sample complexity remains substantially higher than the information-theoretic limit, the latter of which is $nr = O(n)$ in this case since there are only nr free parameters. It is worth noting, however, that all polynomial-time algorithms developed in the literature for tensor completion require a sample size at least exceeding the order of $n^{3/2}$ (Barak and Moitra, 2016). This hints at the (potential) existence of a computational barrier that prevents one from achieving the information-theoretic limit efficiently. Viewed in this light, the spectral method presented herein already achieves near-optimal sample complexity—when restricted to computationally tractable algorithms—if the objective is consistent subspace estimation.

3.9.4 Proof of auxiliary lemmas

Proof of Lemma 3.24. Define the following zero-mean random matrix

$$\mathbf{Z} = p^{-1}\mathbf{A} - \mathbf{A}^*.$$

It is self-evident that

$$p^{-2}(\mathbf{A}\mathbf{A}^\top - \mathbb{E}[\mathbf{A}\mathbf{A}^\top]) = \mathbf{A}^*\mathbf{Z}^\top + \mathbf{Z}\mathbf{A}^{*\top} + (\mathbf{Z}\mathbf{Z}^\top - \mathbb{E}[\mathbf{Z}\mathbf{Z}^\top]),$$

which implies that the identity holds for the off-diagonal part. By the definitions (3.116a) and (3.116b), it follows from the triangle inequality that

$$\begin{aligned} \|\mathbf{M} - \mathbf{M}^*\| &\leq \|\mathcal{P}_{\text{diag}}(\mathbf{A}^*\mathbf{A}^{*\top})\| + 2\|\mathcal{P}_{\text{off-diag}}(\mathbf{A}^*\mathbf{Z}^\top)\| \\ &\quad + \|\mathcal{P}_{\text{off-diag}}(\mathbf{Z}\mathbf{Z}^\top - \mathbb{E}[\mathbf{Z}\mathbf{Z}^\top])\|. \end{aligned} \quad (3.119)$$

In the sequel, we shall discuss how to control the three terms on the right-hand side of (3.119) separately.

Step 1: bounding $\|\mathcal{P}_{\text{diag}}(\mathbf{A}^\mathbf{A}^{*\top})\|$.* It is straightforward to verify that

$$\|\mathcal{P}_{\text{diag}}(\mathbf{A}^*\mathbf{A}^{*\top})\| = \max_{1 \leq l \leq n} \|\mathbf{A}_{l,\cdot}^*\|_2^2 = \|\mathbf{A}^*\|_{2,\infty}^2. \quad (3.120)$$

It thus suffices to bound $\|\mathbf{A}^*\|_{2,\infty}$, which we shall discuss momentarily.

Step 2: bounding $\|\mathcal{P}_{\text{off-diag}}(\mathbf{Z}\mathbf{Z}^\top - \mathbb{E}[\mathbf{Z}\mathbf{Z}^\top])\|$. Define a collection of independent zero-mean random matrices as follows

$$\mathbf{Q}_i := \mathcal{P}_{\text{off-diag}}(\mathbf{Z}_{\cdot,i}\mathbf{Z}_{\cdot,i}^\top), \quad 1 \leq i \leq n^2,$$

with which we can express

$$\mathcal{P}_{\text{off-diag}}(\mathbf{Z}\mathbf{Z}^\top - \mathbb{E}[\mathbf{Z}\mathbf{Z}^\top]) = \sum_i (\mathbf{Q}_i - \mathbb{E}[\mathbf{Q}_i]) = \sum_i \mathbf{Q}_i. \quad (3.121)$$

Here the last relation uses the fact that $\mathcal{P}_{\text{off-diag}}(\mathbb{E}[\mathbf{Z}\mathbf{Z}^\top]) = \sum_i \mathbb{E}[\mathbf{Q}_i] = \mathbf{0}$. Recognizing that the entries of $\mathbf{Z}_{\cdot,i}$ are independently generated, one can see from straightforward calculations that $\mathbb{E}[\mathbf{Q}_i \mathbf{Q}_i^\top]$ is a diagonal matrix, whose diagonal entries satisfy

$$(\mathbb{E}[\mathbf{Q}_i \mathbf{Q}_i^\top])_{l,l} = \mathbb{E}[Z_{l,i}^2] \sum_{j:j \neq l} \mathbb{E}[Z_{j,i}^2] = \frac{1-p}{p} (A_{l,i}^*)^2 \sum_{j:j \neq l} \frac{1-p}{p} (A_{j,i}^*)^2$$

$$\leq \frac{1}{p^2} (A_{l,i}^*)^2 \|\mathbf{A}_{\cdot,i}^*\|_2^2 \leq \frac{1}{p^2} (A_{l,i}^*)^2 \|\mathbf{A}^*\|_{\infty,2}^2$$

for all $1 \leq l \leq n$. Taking into account all samples yields

$$\sum_{i=1}^{n^2} (\mathbb{E}[\mathbf{Q}_i \mathbf{Q}_i^\top])_{l,l} \leq \frac{1}{p^2} \sum_{i=1}^{n^2} (A_{l,i}^*)^2 \|\mathbf{A}^*\|_{\infty,2}^2 \leq \frac{1}{p^2} \|\mathbf{A}^*\|_{2,\infty}^2 \|\mathbf{A}^*\|_{\infty,2}^2$$

for any $1 \leq l \leq n$, which together with the diagonal structure of $\mathbb{E}[\mathbf{Q}_i \mathbf{Q}_i^\top]$ leads to an upper bound on the variance statistic

$$v := \left\| \sum_i \mathbb{E}[\mathbf{Q}_i \mathbf{Q}_i^\top] \right\| = \left| \max_l \left(\sum_i (\mathbb{E}[\mathbf{Q}_i \mathbf{Q}_i^\top])_{l,l} \right) \right| \leq \frac{\|\mathbf{A}^*\|_{2,\infty}^2 \|\mathbf{A}^*\|_{\infty,2}^2}{p^2}.$$

In addition, we identify a suitable truncation level and claim that

$$\mathbb{P}\left\{ \|\mathbf{Q}_i\|_2 \geq L \right\} \leq 2n^{-7} =: q_0, \quad (3.122a)$$

$$\left\| \mathbb{E}[\mathbf{Q}_i \mathbf{1}\{\|\mathbf{Q}_i\|_2 \leq L\}] \right\| \leq 4n^{-7} p^{-2} \|\mathbf{A}^*\|_{\infty,2}^2 =: q_1, \quad (3.122b)$$

where we define $L := 2(4\sqrt{\frac{\log n}{p}} \|\mathbf{A}^*\|_{\infty,2} + \frac{6\log n}{p} \|\mathbf{A}^*\|_\infty)^2$. Armed with these observations, the truncated matrix Bernstein inequality (see Corollary 3.2) taken together with (3.121) reveals that

$$\begin{aligned} \|\mathcal{P}_{\text{off-diag}}(\mathbf{Z}\mathbf{Z}^\top - \mathbb{E}[\mathbf{Z}\mathbf{Z}^\top])\| &\lesssim \sqrt{v \log n} + L \log n + n^2 q_1 \\ &\leq \frac{\sqrt{\log n}}{p} \|\mathbf{A}^*\|_{2,\infty} \|\mathbf{A}^*\|_{\infty,2} + \frac{\log^2 n}{p} \|\mathbf{A}^*\|_{\infty,2}^2 + \frac{\log^3 n}{p^2} \|\mathbf{A}^*\|_\infty^2 \end{aligned} \quad (3.123)$$

with probability $1 - O(n^{-7}) - nq_0 = 1 - O(n^{-7})$, provided that $p \gtrsim n^{-5}$.

Step 3: bounding $\|\mathcal{P}_{\text{off-diag}}(\mathbf{A}^ \mathbf{Z}^\top)\|$.* This term can be controlled in a similar fashion as $\|\mathcal{P}_{\text{off-diag}}(\mathbf{Z}\mathbf{Z}^\top - \mathbb{E}[\mathbf{Z}\mathbf{Z}^\top])\|$. We thus omit the details and only state the result as follows:

$$\begin{aligned} \|\mathcal{P}_{\text{off-diag}}(\mathbf{A}^* \mathbf{Z}^\top)\| &\lesssim \|\mathbf{A}^*\|_{\infty,2} \|\mathbf{A}^*\| \sqrt{\frac{\log n}{p}} \\ &\quad + \sqrt{\frac{\log^3 n}{p}} \|\mathbf{A}^*\|_{\infty,2}^2 + \|\mathbf{A}^*\|_{\infty,2} \|\mathbf{A}^*\|_\infty \frac{\log^2 n}{p} \end{aligned} \quad (3.124)$$

holds with probability at least $1 - O(n^{-7})$.

Step 4: To finish up, we are in need of bounding $\|\mathbf{A}^*\|_{\infty,2}$, $\|\mathbf{A}^*\|_{2,\infty}$ and $\|\mathbf{A}^*\|_\infty$, which is accomplished in the following lemma.

Lemma 3.27. Suppose that $\mu r^2 \leq n$. Then one has

$$\|\mathbf{A}^*\|_\infty \leq \frac{\mu^{3/2} r \nu_{\max}}{n^{3/2}}, \quad \|\mathbf{A}^*\|_{\infty,2} \leq \frac{\mu \sqrt{2r} \nu_{\max}}{n}, \quad \|\mathbf{A}^*\|_{2,\infty} \leq \sqrt{\frac{2\mu r}{n}} \nu_{\max}.$$

Taking Lemma 3.27 collectively with (3.120), (3.123), (3.124) and combining terms, we arrive at

$$(3.120) + (3.123) + (3.124) \lesssim \left(\frac{\mu^{3/2} r \sqrt{\log n}}{n^{3/2} p} + \sqrt{\frac{\mu^2 r \log n}{n^2 p}} + \frac{\mu r}{n} \right) \nu_{\max}^2$$

with probability at least $1 - O(n^{-7})$, provided that $\mu \log n \leq n$ and $p \gtrsim \frac{\mu^{3/2} r \log^{2.5} n}{n^{3/2}}$. This taken together with (3.119) concludes the proof.

Proof of the relation (3.122). We first make note of a connection between \mathbf{Q}_i and $\mathbf{Z}_{\cdot,i}$ as follows

$$\|\mathbf{Q}_i\| \leq \|\mathbf{Z}_{\cdot,i} \mathbf{Z}_{\cdot,i}^\top\| + \|\mathcal{P}_{\text{diag}}(\mathbf{Z}_{\cdot,i} \mathbf{Z}_{\cdot,i}^\top)\| \leq 2\|\mathbf{Z}_{\cdot,i}\|_2^2, \quad (3.125)$$

which motivates us to first control the size of $\mathbf{Z}_{\cdot,i}$. By construction, each entry $Z_{j,i}$ can be written as $Z_{j,i} = (\frac{1}{p} \delta_{j,i} - 1) A_{j,i}^*$, where $\{\delta_{j,i}\}$ is a collection of independent Bernoulli random variables with mean p . This observation allows one to derive

$$\begin{aligned} B_z &:= \max_{i,j} |Z_{j,i}| \leq \frac{1}{p} \|\mathbf{A}^*\|_\infty; \\ v_z &:= \mathbb{E}[\|\mathbf{Z}_{\cdot,i}\|_2^2] = \frac{1-p}{p} \sum_j (A_{j,i}^*)^2 \leq \frac{1}{p} \|\mathbf{A}^*\|_{\infty,2}^2. \end{aligned}$$

The matrix Bernstein inequality (see Corollary 3.3) then yields

$$\begin{aligned} \|\mathbf{Z}_{\cdot,i}\|_2 &\leq 4\sqrt{v_z \log n} + 6B_z \log n \\ &\leq \left(4\sqrt{\frac{\log n}{p}} \|\mathbf{A}^*\|_{\infty,2} + \frac{6 \log n}{p} \|\mathbf{A}^*\|_\infty \right) =: \beta_z \end{aligned} \quad (3.126)$$

with probability at least $1 - 2n^{-7}$, which combined with (3.125) gives

$$\mathbb{P}\left\{\|\mathbf{Q}_i\|_2 \geq 2\beta_z^2\right\} \leq 2n^{-7}. \quad (3.127)$$

Recalling that $\mathbb{E}[\mathbf{Q}_i] = \mathbf{0}$, one can derive

$$\begin{aligned} \left\|\mathbb{E}[\mathbf{Q}_i \mathbf{1}\{\|\mathbf{Q}_i\|_2 \leq 2\beta_z^2\}]\right\| &= \left\|\mathbb{E}[\mathbf{Q}_i] - \mathbb{E}[\mathbf{Q}_i \mathbf{1}\{\|\mathbf{Q}_i\|_2 > 2\beta_z^2\}]\right\| \\ &= \left\|\mathbb{E}[\mathbf{Q}_i \mathbf{1}\{\|\mathbf{Q}_i\|_2 > 2\beta_z^2\}]\right\| \stackrel{(i)}{\leq} \mathbb{P}\{\|\mathbf{Q}_i\|_2 > 2\beta_z^2\} \cdot \frac{2}{p^2} \|\mathbf{A}_{\cdot,i}^*\|_2^2 \\ &\stackrel{(ii)}{\leq} \frac{4}{n^7 p^2} \|\mathbf{A}^*\|_{\infty,2}^2. \end{aligned}$$

Here, (i) relies on (3.125) and the fact $\|\mathbf{Z}_{\cdot,i}\|_2 \leq p^{-1} \|\mathbf{A}_{\cdot,i}^*\|_2$ (by construction), whereas (ii) results from the calculation in (3.126).

Proof of Lemma 3.25. Define the normalized tensor factors as

$$\bar{\mathbf{w}}_i^* := \mathbf{w}_i^* / \|\mathbf{w}_i^*\|_2 \quad (1 \leq i \leq r),$$

and it is convenient to introduce the following auxiliary matrices that contain information about them:

$$\bar{\mathbf{W}}^* := [\bar{\mathbf{w}}_1^*, \dots, \bar{\mathbf{w}}_r^*], \quad \bar{\mathbf{W}}_{\text{lift}}^* := [\bar{\mathbf{w}}_1^* \otimes \bar{\mathbf{w}}_1^*, \dots, \bar{\mathbf{w}}_r^* \otimes \bar{\mathbf{w}}_r^*].$$

Additionally, we introduce a diagonal matrix $\mathbf{D}^* \in \mathbb{R}^{r \times r}$ whose diagonal entries are given by

$$[\mathbf{D}^*]_{i,i} = \|\mathbf{w}_i^*\|_2^3 = v_i, \quad 1 \leq i \leq r.$$

The matrices introduced above allow one to express $\mathbf{A}^* = \bar{\mathbf{W}}^* \mathbf{D}^* \bar{\mathbf{W}}_{\text{lift}}^{*\top}$ and $\mathbf{A}^* \mathbf{A}^{*\top}$ as follows

$$\mathbf{A}^* \mathbf{A}^{*\top} = \bar{\mathbf{W}}^* \mathbf{D}^* \bar{\mathbf{W}}_{\text{lift}}^{*\top} \bar{\mathbf{W}}_{\text{lift}}^* \mathbf{D}^* \bar{\mathbf{W}}^* \mathbf{A}^{*\top}. \quad (3.128)$$

Clearly, the rank of $\mathbf{A}^* \mathbf{A}^{*\top}$ is bounded above by r , and hence it suffices to lower bound $\lambda_i(\mathbf{A}^* \mathbf{A}^{*\top})$ when $i \leq r$.

In order to characterize the spectrum of $\mathbf{A}^* \mathbf{A}^{*\top}$, we first look at the eigenvalues of $\bar{\mathbf{W}}^{*\top} \bar{\mathbf{W}}^*$ and $\bar{\mathbf{W}}_{\text{lift}}^{*\top} \bar{\mathbf{W}}_{\text{lift}}^*$. Write

$$\bar{\mathbf{W}}^{*\top} \bar{\mathbf{W}}^* = \mathbf{I}_r + \mathbf{R}, \quad \text{and} \quad \bar{\mathbf{W}}_{\text{lift}}^{*\top} \bar{\mathbf{W}}_{\text{lift}}^* = \mathbf{I}_r + \mathbf{R}_{\text{lift}} \quad (3.129)$$

for some residual matrices $\mathbf{R}, \mathbf{R}_{\text{lift}} \in \mathbb{R}^{r \times r}$ (which are off-diagonal matrices). By virtue of the definition (3.112), we immediately obtain

$$\|\mathbf{R}\|_{\infty} \leq \sqrt{\mu/n}, \quad \text{and} \quad \|\mathbf{R}_{\text{lift}}\|_{\infty} \leq \mu/n,$$

thus indicating that

$$\|\mathbf{R}\| \leq r \|\mathbf{R}\|_{\infty} \leq r \sqrt{\mu/n}, \quad \|\mathbf{R}_{\text{lift}}\| \leq r \|\mathbf{R}_{\text{lift}}\|_{\infty} \leq \mu r/n. \quad (3.130)$$

Putting these together with (3.129) and invoking Weyl's inequality give

$$\max_i \left| \lambda_i(\bar{\mathbf{W}}^{\star\top} \bar{\mathbf{W}}^{\star}) - 1 \right| \leq \|\mathbf{R}\| \leq r \sqrt{\mu/n}, \quad (3.131)$$

which together with the assumption $\mu r^2 \leq n$ further reveals that

$$\|\bar{\mathbf{W}}^{\star}\| = \sqrt{\lambda_1(\bar{\mathbf{W}}^{\star\top} \bar{\mathbf{W}}^{\star})} \leq \sqrt{1 + r \sqrt{\mu/n}} \leq 2. \quad (3.132)$$

We now return to study $\mathbf{A}^{\star} \mathbf{A}^{\star\top}$. In view of (3.128) and (3.129), one can decompose $\mathbf{A}^{\star} \mathbf{A}^{\star\top}$ into the following two terms

$$\mathbf{A}^{\star} \mathbf{A}^{\star\top} = \underbrace{\bar{\mathbf{W}}^{\star} (\mathbf{D}^{\star})^2 \bar{\mathbf{W}}^{\star\top}}_{=: \mathbf{G}_1} + \underbrace{\bar{\mathbf{W}}^{\star} \mathbf{D}^{\star} \mathbf{R}_{\text{lift}} \mathbf{D}^{\star} \bar{\mathbf{W}}^{\star\top}}_{=: \mathbf{G}_2}. \quad (3.133)$$

Making use of the bounds (3.130) and (3.132) immediately leads to

$$\|\mathbf{G}_2\| \leq \|\bar{\mathbf{W}}^{\star}\|^2 \|\mathbf{D}^{\star}\|^2 \|\mathbf{R}_{\text{lift}}\| \leq 4\mu r \nu_{\max}^2 / n.$$

Regarding \mathbf{G}_1 , it can be directly seen that the non-zero eigenvalues of \mathbf{G}_1 coincide with those of $\mathbf{D}^{\star} \bar{\mathbf{W}}^{\star\top} \bar{\mathbf{W}}^{\star} \mathbf{D}^{\star}$, where the latter can be decomposed into

$$\mathbf{D}^{\star} \bar{\mathbf{W}}^{\star\top} \bar{\mathbf{W}}^{\star} \mathbf{D}^{\star} = (\mathbf{D}^{\star})^2 + \mathbf{D}^{\star} \mathbf{R} \mathbf{D}^{\star}.$$

As a result, for any $1 \leq i \leq r$ one can derive

$$\begin{aligned} \left| \lambda_i(\mathbf{G}_1) - \lambda_i((\mathbf{D}^{\star})^2) \right| &= \left| \lambda_i(\mathbf{D}^{\star} \bar{\mathbf{W}}^{\star\top} \bar{\mathbf{W}}^{\star} \mathbf{D}^{\star}) - \lambda_i((\mathbf{D}^{\star})^2) \right| \\ &\leq \|\mathbf{D}^{\star} \mathbf{R} \mathbf{D}^{\star}\| \leq \|\mathbf{D}^{\star}\|^2 \|\mathbf{R}\| \leq r \sqrt{\frac{\mu}{n}} \nu_{\max}^2. \end{aligned}$$

This taken together with the decomposition (3.133) leads to

$$\left| \lambda_i(\mathbf{A}^{\star} \mathbf{A}^{\star\top}) - \lambda_i(\mathbf{G}_1) \right| \leq \|\mathbf{G}_2\| \leq \frac{4\mu r \nu_{\max}^2}{n},$$

thus indicating that

$$\begin{aligned} & \left| \lambda_i(\mathbf{A}^* \mathbf{A}^{*\top}) - \lambda_i((\mathbf{D}^*)^2) \right| \\ & \leq \left| \lambda_i(\mathbf{G}_1) - \lambda_i((\mathbf{D}^*)^2) \right| + \left| \lambda_i(\mathbf{A}^* \mathbf{A}^{*\top}) - \lambda_i(\mathbf{G}_1) \right| \\ & \leq r \sqrt{\frac{\mu}{n}} \nu_{\max}^2 + \frac{4\mu r \nu_{\max}^2}{n} \leq 8 \max \left\{ \sqrt{\frac{\mu}{n}}, \frac{\mu}{n} \right\} r \nu_{\max}^2. \end{aligned}$$

If $16 \max \left\{ \frac{\mu}{n}, \sqrt{\frac{\mu}{n}} \right\} r \nu_{\max}^2 \leq \nu_{\min}^2$, then one has $|\lambda_i(\mathbf{A}^* \mathbf{A}^{*\top}) - \lambda_i((\mathbf{D}^*)^2)| \leq \nu_{\min}^2/2$. In addition, letting $\nu_{(i)}$ be the i -th largest element in $\{\nu_i\}_{1 \leq i \leq r}$, we have $\lambda_i((\mathbf{D}^*)^2) = \nu_{(i)}^2$ and hence arrive at

$$\nu_{\min}^2/2 \leq \nu_{(i)}^2 - \nu_{\min}^2/2 \leq \lambda_i(\mathbf{A}^* \mathbf{A}^{*\top}) \leq \nu_{(i)}^2 + \nu_{\min}^2/2 \leq 2\nu_{\max}^2$$

for any $1 \leq i \leq r$, as claimed.

Proof of Lemma 3.27. Define the following two matrices containing information about the tensor factors:

$$\mathbf{W}^* := [\mathbf{w}_1^*, \dots, \mathbf{w}_r^*] \in \mathbb{R}^{n \times r}, \quad (3.134a)$$

$$\mathbf{W}_{\text{lift}}^* := [\mathbf{w}_1^* \otimes \mathbf{w}_1^*, \dots, \mathbf{w}_r^* \otimes \mathbf{w}_r^*] \in \mathbb{R}^{n^2 \times r}. \quad (3.134b)$$

Given that $\mathbf{W}^{*\top} \mathbf{W}^* = [\langle \mathbf{w}_i^*, \mathbf{w}_j^* \rangle]_{1 \leq i, j \leq r}$, its diagonal part satisfies

$$\left\| \mathcal{P}_{\text{diag}}(\mathbf{W}^{*\top} \mathbf{W}^*) \right\| = \left\| \text{diag} \left([\|\mathbf{w}_i^*\|_2^2]_{1 \leq i \leq r} \right) \right\| \leq \nu_{\max}^{2/3}.$$

In addition, the off-diagonal part of $\mathbf{W}^{*\top} \mathbf{W}^*$ satisfies

$$\begin{aligned} \left\| \mathcal{P}_{\text{off-diag}}(\mathbf{W}^{*\top} \mathbf{W}^*) \right\| & \leq \sqrt{\sum_{i \neq j} |\langle \mathbf{w}_i^*, \mathbf{w}_j^* \rangle|^2} \leq \sqrt{\frac{\mu}{n} \sum_{i \neq j} \|\mathbf{w}_i^*\|_2^2 \|\mathbf{w}_j^*\|_2^2} \\ & \leq \sqrt{\frac{\mu r^2}{n}} \max_i \|\mathbf{w}_i^*\|_2^2 = \nu_{\max}^{2/3} \sqrt{\frac{\mu r^2}{n}}, \end{aligned}$$

where the second inequality relies on the definition (3.112) of the incoherence parameter, and the last relation follows from the definition of ν_{\max} in (3.111). Consequently, if $\mu r^2 \leq n$, then

$$\|\mathbf{W}^*\|^2 = \|\mathbf{W}^{*\top} \mathbf{W}^*\| \leq \left\| \mathcal{P}_{\text{diag}}(\mathbf{W}^{*\top} \mathbf{W}^*) \right\| + \left\| \mathcal{P}_{\text{off-diag}}(\mathbf{W}^{*\top} \mathbf{W}^*) \right\|$$

$$\leq \nu_{\max}^{2/3} + \nu_{\max}^{2/3} \sqrt{\frac{\mu r^2}{n}} \leq 2\nu_{\max}^{2/3}. \quad (3.135)$$

Repeating similar arguments also reveals that

$$\|\mathbf{W}_{\text{lift}}^*\|^2 \leq 2\nu_{\max}^{4/3}. \quad (3.136)$$

Next, it is readily seen from the definition (3.112) that

$$\begin{aligned} \|\mathbf{W}^*\|_{2,\infty} &\leq \sqrt{r} \max_i \|\mathbf{w}_i^*\|_\infty \leq \sqrt{\frac{\mu r}{n}} \max_i \|\mathbf{w}_i^*\|_2 = \sqrt{\frac{\mu r}{n}} \nu_{\max}^{1/3}, \\ \|\mathbf{W}_{\text{lift}}^*\|_{2,\infty} &\leq \sqrt{r} \max_i \|\mathbf{w}_i^*\|_\infty^2 \leq \frac{\mu \sqrt{r}}{n} \max_i \|\mathbf{w}_i^*\|_2^2 = \frac{\mu \sqrt{r}}{n} \nu_{\max}^{2/3}. \end{aligned}$$

Combining these bounds with (3.135) and (3.136) immediately yields

$$\begin{aligned} \|\mathbf{A}^*\|_{\infty,2} &= \left\| \mathbf{W}^* (\mathbf{W}_{\text{lift}}^*)^\top \right\|_{\infty,2} \leq \|\mathbf{W}^*\| \|\mathbf{W}_{\text{lift}}^*\|_{2,\infty} \leq \frac{\mu \sqrt{2r}}{n} \nu_{\max}, \\ \|\mathbf{A}^*\|_\infty &= \left\| \mathbf{W}^* (\mathbf{W}_{\text{lift}}^*)^\top \right\|_\infty \leq \|\mathbf{W}^*\|_{2,\infty} \|\mathbf{W}_{\text{lift}}^*\|_{2,\infty} \leq \frac{\mu^{3/2} r}{n^{3/2}} \nu_{\max}, \\ \|\mathbf{A}^*\|_{2,\infty} &= \left\| \mathbf{W}^* (\mathbf{W}_{\text{lift}}^*)^\top \right\|_{2,\infty} \leq \|\mathbf{W}^*\|_{2,\infty} \|\mathbf{W}_{\text{lift}}^*\| \leq \sqrt{\frac{2\mu r}{n}} \nu_{\max}. \end{aligned}$$

3.10 Notes

This section provides further pointers to the applications studied in this chapter, and singles out a brief list of applications we have omitted.

Before proceeding, it is worth pointing out several important facts. First, for many applications (e.g., phase retrieval, matrix and tensor completion), the spectral method alone does not allow for perfect reconstruction of the unknowns even when it is information-theoretically feasible to do so. Instead, the spectral method frequently serves as a suitable initialization step for these applications, and its estimate can often be further refined by means of nonconvex optimization algorithms like gradient descent and alternating minimization; see Chi *et al.* (2019) and Jain and Kar (2017) for overviews of recent advances. Second, throughout this chapter, we have assumed that the underlying matrix is exactly low-rank, and in addition the spectral methods deployed know the correct rank. However, in reality, data matrices are rarely exactly

low-rank. It is therefore of great importance to develop and analyze methods that can handle such misspecified cases. When the reconstruction error of the matrix is considered, several methods are capable of achieving graceful tradeoff between the estimation error and the approximation error, without knowing the correct rank, e.g. e.g., Koltchinskii *et al.* (2011), Chatterjee (2014), and Negahban and Wainwright (2011). In addition, further discussions (e.g., convex relaxation approaches and nonconvex landscape analysis) about several of these applications can be found in Candès (2014), Wainwright (2019), Wright and Ma (2020), and Zhang *et al.* (2020d).

PCA, factor models and covariance estimation. PCA and factor models are among the most classic and extensively studied topics in statistics (Anderson, 1962; Fan *et al.*, 2020b; Wainwright, 2019). The model considered in Section 3.3.1 has been studied by, for example, Johnstone (2001), Paul (2007), Nadler (2008), Perry *et al.* (2018), Xie *et al.* (2020), Wang and Fan (2017), Fan *et al.* (2018c), and Bao *et al.* (2020) under the name of spiked covariance models, covering both the finite-sample regime and high-dimensional asymptotics. A more recent strand of work extended the theory to accommodate heteroskedastic noise and missing data (including heterogeneous missing patterns) (Lounici, 2014; Zhang *et al.*, 2021; Cai *et al.*, 2021a; Zhu *et al.*, 2019), as well as exponential family distributions (Liu *et al.*, 2018). In addition to providing the distance control between the spectral estimate and the true principle subspace, Koltchinskii and Lounici (2016) and Fan *et al.* (2019b) also studied the bias of the spectral estimate under various types of data distributions. It is clearly impossible to review the enormous literature in a monograph of this length; the interested reader is referred to the overview papers Johnstone and Paul (2018), Fan *et al.* (2018a), Vaswani *et al.* (2018), and Balzano *et al.* (2018) and the recent books Fan *et al.* (2020b) and Wainwright (2019) for overviews of contemporary developments on this topic (with particular emphasis on high-dimensional data). In addition, this monograph does not account for the sparsity structure, or a superposition of low-rank and sparsity structure, where are commonly imposed on either the covariance matrix or the precision matrix (Johnstone and Lu, 2009; Ma, 2013; Vu and Lei,

2012; Cai *et al.*, 2013; Candès *et al.*, 2011; Chandrasekaran *et al.*, 2011; Chandrasekaran *et al.*, 2010). These additional structural assumptions play a crucial role in further dimension reduction and are useful for, say, learning graphical models, video surveillance in computer vision, and portfolio allocation and risk managements in finance; see Fan *et al.* (2020b), Vidal *et al.* (2016), Wainwright (2019), and Wright and Ma (2020) for more detailed discussions.

Applications of PCA in statistical and econometric modeling. PCA has been widely applied to estimate dimension-reduced spaces in multiple-index models (Li, 1992; Duan and Li, 1991; Cook, 2007; Xia *et al.*, 2002; Li, 2018), and latent factors in econometric modeling (Forni *et al.*, 2000; Stock and Watson, 2002; Bai and Ng, 2002; Bai, 2003; Bai, 2009; Ahn and Horenstein, 2013; Fan *et al.*, 2015; Fan *et al.*, 2016). For recent reviews of this topic, we refer the readers to Stock and Watson (2016) for dynamic factor models with applications to macroeconomics, to Bai and Wang (2016) for time series and panel data models, to Fan *et al.* (2021b) for robust factor models and large covariance estimation, and to Fan *et al.* (2021a) for factor models and their broader applications to econometric learning. In particular, factor models have been frequently employed to adjust correlated covariates in high-dimensional model selection, large-scale inference, predictions, treatment evaluations, among others; see Fan *et al.* (2021b) and Fan *et al.* (2021a) and the references therein.

Graph clustering and community recovery. Spectral methods—possibly coupled with other subsequent refining schemes like k -means—are among the most widely used algorithms for graph clustering (McSherry, 2001; Rohe *et al.*, 2011; Balakrishnan *et al.*, 2011; Chaudhuri *et al.*, 2012; Fishkind *et al.*, 2013; Sarkar and Bickel, 2015; Jin, 2015; Gao *et al.*, 2017; Zhang and Zhou, 2020; Newman, 2013; Chen and Hero, 2015; Zhang *et al.*, 2020c; Le and Levina, 2015; Jin, 2015; Le *et al.*, 2018; Chen *et al.*, 2020c). While a large fraction of earlier papers required the average vertex degree to be significantly larger than $\log n$, Lei and Rinaldo (2015) broadened the coverage of the theory by accommodating sparse graphs with average degrees as low as $O(\log n)$. This, however,

should be differentiated from the ultra-sparse regime with average degrees $O(1)$; in this scenario, spectral methods based on vanilla adjacency matrices no longer work, and more intelligent designs are needed to effectively detect the communities (Coja-Oghlan, 2010; Massoulié, 2014; Chin *et al.*, 2015; Le *et al.*, 2017). The theory available for spectral clustering extends far beyond the two-community SBM presented herein, examples including SBMs with growing communities (Rohe *et al.*, 2011), degree-corrected SBMs (Lei and Rinaldo, 2015; Lei and Zhu, 2014), graphs with locality (Chen *et al.*, 2016a), mixed membership models Fan *et al.* (2019a) and Han *et al.* (2019), hyper-graphs (Ahn *et al.*, 2018; Michoel and Nachtergael, 2012; Cole and Zhu, 2020), and directed graphs (Wang *et al.*, 2020). An abundance of other paradigms, most notably convex relaxation, have also proved effective for clustering (Jalali *et al.*, 2011; Amini *et al.*, 2013; Abbe *et al.*, 2016; Hajek *et al.*, 2016; Cai and Li, 2015; Zhao *et al.*, 2012; Li *et al.*, 2021b; Zhang and Zhou, 2020; Yuan and Qu, 2018; Fei and Chen, 2018a; Fei and Chen, 2019). We recommend the article Abbe (2017) for an overview of recent developments.

Gaussian mixture models. The Gaussian mixture model is among the most classic and convenient statistical models to capture the effect of multi-modal and heterogeneous data (e.g., Pearson (1894), Titterington *et al.* (1985), Xu and Jordan (1996), Dasgupta (1999), Hsu and Kakade (2013), Kalai *et al.* (2010), Balakrishnan *et al.* (2017), Xu *et al.* (2016), Jin *et al.* (2016b), Fei and Chen (2018b), Jin *et al.* (2017), Dan *et al.* (2020), and Han *et al.* (2021)). Unlike parameter estimation (e.g., estimating the centers) that does not require center separation (Wu and Yang, 2020), the feasibility of reliable clustering in Gaussian mixture models is dictated by the minimum center separation (Lu and Zhou, 2016; Cai and Zhang, 2018; Ndaoud, 2018; Giraud and Verzelen, 2019; Chen and Yang, 2020). While spectral methods naturally come into mind for the clustering task and have been frequently applied in the literature (Von Luxburg, 2007; Vempala and Wang, 2004; Kannan *et al.*, 2008; Kumar and Kannan, 2010; Awasthi and Sheffet, 2012), sharp statistical analysis of spectral clustering (and its variants) has been lacking until recently (Ndaoud, 2018; Löffler *et al.*, 2021; Srivastava

et al., 2019; Abbe *et al.*, 2020a). While it might be tempting to impose a minimum spectral gap requirement on the matrix Θ^* (cf. (3.41)) in order to invoke the $\sin \Theta$ theorems, such a condition can be dropped as long as an appropriate spectral clustering scheme is employed (Löffler *et al.*, 2021). Encouragingly, spectral clustering (with the aid of k -means) also achieves information-theoretically optimal mis-clustering rate exponents for a couple of scenarios (Löffler *et al.*, 2021; Abbe *et al.*, 2020a).

Ranking from pairwise comparisons. Deploying spectral methods to address ranking tasks has a long history, dating back at least to Seeley (1949). We refer the readers to Vigna (2016) for a historical account of this subject. The specific instance of spectral methods introduced here was due to Negahban *et al.* (2016), and has been subsequently analyzed in multiple papers (Rajkumar and Agarwal, 2014; Chen and Suh, 2015; Jang *et al.*, 2016; Chen *et al.*, 2020b). It bears close similarity to the celebrated PageRank algorithm heavily used by Google (Page *et al.*, 1999). Negahban *et al.* (2016) developed the first ℓ_2 statistical guarantees when estimating the underlying score vector, accounting for missing data and general comparison graphs as well. The ℓ_2 guarantees for random comparison graphs were further sharpened in Chen and Suh (2015) (which closed the logarithmic gap). Note, however, that the ℓ_2 score estimation error bounds alone typically do not imply the ranking accuracy. Motivated by this inadequacy, Chen and Suh (2015) directly studied the top- K ranking accuracy, by demonstrating the optimality of spectral ranking followed by an iterative refinement scheme. However, this left open another question regarding whether the follow-up refinement step is necessary in achieving optimal ranking accuracy. Jang *et al.* (2016) attempted to address this question by establishing desired ranking accuracy of spectral methods when the number of pairwise comparisons available is large. A complete picture was subsequently obtained by Chen *et al.* (2019b), which proved the optimality of spectral methods in top- K ranking all the way to the sample-starved regime. Moving beyond exact top- K ranking, the recent work Chen *et al.* (2020b) studied the capability (and limitations) of spectral methods in handling partial recovery of the top- K ranked items. Moving beyond the BTL model, there are also a number of other ranking

models that have been extensively studied in the literature (e.g., the Plackett-Luce model for multi-way comparisons (Hunter, 2004; Hajek *et al.*, 2014; Oh and Shah, 2014; Agarwal *et al.*, 2018), the stochastically transitive model (Shah *et al.*, 2016; Shah *et al.*, 2019)), which are beyond the scope of the present monograph.

Phase retrieval. Netrapalli *et al.* (2015) proposed the first spectral method (cf. Section 3.7.2) for phase retrieval, and established the performance guarantees when the sample size exceeds $m \gtrsim n \log^3 n$. The theoretical support was then tightened by Candès *et al.* (2015a), allowing the sample size to be as low as $m \asymp n \log n$. Similar theory was provided for the random coded diffraction pattern model in Candès *et al.* (2015a). Several variations and generalizations of the spectral method have been further proposed to improve performance. The first order-wise optimal spectral method for phase retrieval was proposed by Chen and Candès (2017), based on the truncation idea. This method has multiple variants (Zhang *et al.*, 2016; Li *et al.*, 2020c; Wang *et al.*, 2018a), and has been shown to be robust against noise and corruptions. The precise asymptotic characterization of the spectral method was first obtained in Lu and Li (2020). Based on this characterization, Mondelli and Montanari (2019) and Luo *et al.* (2019) later devised optimal designs of spectral methods in phase retrieval when the sensing matrix follows the Gaussian design, where its sensitivity to model mismatch was studied in Monardo and Chi (2019). Ma *et al.* (2021b) and Dudeja *et al.* (2020) explored similar questions when the sensing matrix is Haar distributed (e.g., an isotropically random unitary matrix). The spectral method presented herein has been used to seed a follow-up procedure that in turn enhances estimation accuracy; see, e.g., Netrapalli *et al.* (2015), Candès *et al.* (2015a), Sanghavi *et al.* (2017), Goldstein and Studer (2018), Bahmani and Romberg (2017), Ma *et al.* (2020), Chandra *et al.* (2019), Dhifallah *et al.* (2017), Qu *et al.* (2019), Zhang *et al.* (2017), Ma *et al.* (2019), Tan and Vershynin (2019), Jeong and Güntürk (2017), Cai *et al.* (2019b), and Salehi *et al.* (2018a). An alternative to the spectral method, based on a nullspace approach, has been proposed in Chen *et al.* (2017). Fannjiang and Strohmer (2020) provided an extensive discussion on initialization strategies for algorithmic phase retrieval,

including but not limited to various forms of spectral methods. Sparse phase retrieval is another important topic when the signal of interest is assumed to be a sparse vector; we refer the interested reader to Li and Voroninski (2013), Oymak *et al.* (2015), Chen *et al.* (2015b), Cai *et al.* (2016), Wang *et al.* (2018b), Jagatap and Hegde (2019), Yang *et al.* (2019), Soltanolkotabi (2019), Yang *et al.* (2016), Zhang *et al.* (2018b), Salehi *et al.* (2018b), Shechtman *et al.* (2014), Yuan *et al.* (2019), and Eldar and Mendelson (2014) and additional references cited therein.

Matrix completion. Regarding matrix completion, the spectral method was originally proposed in Achlioptas and McSherry (2007) and Keshavan *et al.* (2010) to estimate (approximately) low-rank matrices in the face of missing data and random corruptions. Similar to phase retrieval, the estimate returned by the spectral method is employed as a suitable initialization to enable fast convergence of nonconvex iterative procedures; see, e.g., Keshavan *et al.* (2010), Keshavan *et al.* (2009), Jain *et al.* (2013), Hardt (2014), Sun and Luo (2016), Chen and Wainwright (2015), Zheng and Lafferty (2016), Boumal and Absil (2015), Wei *et al.* (2016), Chen *et al.* (2020a), Ma *et al.* (2020), Zhang *et al.* (2018c), Jin *et al.* (2016a), and Charisopoulos *et al.* (2021). Moreover, there are several nuclear norm penalized estimators that also bear close relevance to the spectral method, e.g., Koltchinskii *et al.* (2011). We also remark in passing that there are other estimators that can effectively handle the case when the underlying matrix is not exactly low-rank, including but not limited to Universal Singular Value Thresholding (Chatterjee, 2014) and its soft-thresholded version (Koltchinskii *et al.*, 2011). In addition, while our discussion focuses on clean data and uniform random sampling patterns, it is of great importance to study various noisy and quantized scenarios (Keshavan *et al.*, 2009; Candès and Plan, 2010; Cao and Xie, 2015; Klopp, 2014; Chen and Wainwright, 2015; McRae and Davenport, 2021; Davenport *et al.*, 2014; Ma *et al.*, 2020; Zhang *et al.*, 2018d; Krahmer and Stöger, 2021), as well as non-uniform or deterministic sampling patterns (Foucart *et al.*, 2020; Negahban and Wainwright, 2012; Shapiro *et al.*, 2018).

Tensor completion and estimation. Unfolding-based spectral methods have been frequently adopted to deal with various tensor estimation problems including tensor PCA, tensor decomposition, tensor completion, and so on (Richard and Montanari, 2014; Montanari and Sun, 2018; Han *et al.*, 2020; Zhang and Xia, 2018; Cai *et al.*, 2021a; Cai *et al.*, 2021b; Xia and Yuan, 2019; Xia *et al.*, 2021; Moitra and Wein, 2019; Liu and Moitra, 2020; Zhang *et al.*, 2020a; Xia *et al.*, 2020; Tong *et al.*, 2021c). When it comes to tensor completion, the first near-optimal ℓ_2 statistical analysis of spectral methods was due to Montanari and Sun (2018), which was subsequently extended by Cai *et al.* (2021a) to enable $\ell_{2,\infty}$ error control. The readers interested in higher-order tensors (beyond third-order tensors) can consult Montanari and Sun (2018) and Richard and Montanari (2014). In addition, the theory and algorithm presented herein focus attention on the regime where $r < n$, and fall short of accommodating “over-complete” tensors when r rises above n . Certain “contraction” tricks are needed in order to cope with the over-complete regime; see Hopkins *et al.* (2016) and Montanari and Sun (2018).

An extensive but non-exhaustive list of other applications. Finally, the list of applications discussed in this monograph is clearly far from comprehensive. Spectral methods have been successfully applied to a plethora of other problems, including but not limited to the following topics:

- matrix sensing: Tu *et al.* (2016), Zheng and Lafferty (2015), Ma *et al.* (2021a), Chen *et al.* (2021a), Tong *et al.* (2021b), Tong *et al.* (2021a), and Lee *et al.* (2017);
- phase synchronization and group synchronization: Singer (2011), Abbe *et al.* (2020b), and Ling (2020);
- joint matching and map synchronization: Chen *et al.* (2014), Pachauri *et al.* (2013), Shen *et al.* (2016), Bajaj *et al.* (2018), Sun *et al.* (2018), Sun *et al.* (2019), Huang *et al.* (2019a), and Huang *et al.* (2019b);

- covariance sketching and quadratic sensing: Li *et al.* (2021c), Charisopoulos *et al.* (2021), Sanghavi *et al.* (2017), and Chi and Fu (2017);
- blind deconvolution and blind calibration: Li *et al.* (2019), Ma *et al.* (2020), Huang and Hand (2018), Charisopoulos *et al.* (2019), Chen *et al.* (2021d), Li *et al.* (2018), and Cambareri and Jacques (2016);
- blind demixing: Ling and Strohmer (2019) and Dong and Shi (2018);
- low-rank phase retrieval and phaseless PCA: Vaswani *et al.* (2017), Nayer *et al.* (2019), and Vaswani (2020);
- canonical correlation analysis (CCA): Cai and Zhang (2018) and Ge *et al.* (2016);
- sparse PCA: Amini and Wainwright (2008) and Johnstone and Lu (2009);
- mixed linear regression: Yi *et al.* (2014), Ghosh and Kannan (2020), and Kwon *et al.* (2021);
- finding hidden cliques: Alon *et al.* (1998);
- joint image alignment: Chen and Candès (2018);
- robust subspace recovery and robust PCA: Yi *et al.* (2016), Ntrapalli *et al.* (2014), Cherapanamjeri *et al.* (2017), Tong *et al.* (2021b), Zhu *et al.* (2018), and Maunu *et al.* (2019);
- contextual stochastic block models: Binkiewicz *et al.* (2017) and Abbe *et al.* (2020a);
- learning neural networks: Zhong *et al.* (2017) and Fu *et al.* (2020);
- topic modeling: Ke and Wang (2017);
- crowd sourcing: Ghosh *et al.* (2011), Dalvi *et al.* (2013), Karger *et al.* (2013), Zhang *et al.* (2014), and Karger *et al.* (2014);
- meta learning: Kong *et al.* (2020), Du *et al.* (2020), and Tripuraneni *et al.* (2021);

- subspace clustering: Eriksson *et al.* (2012) and Li and Gu (2021);
- state aggregation and compression of Markov chains: Zhang and Wang (2020) and Duan *et al.* (2019);
- causal inference: Amjad *et al.* (2018);
- passive imaging: Lee *et al.* (2018).

For the sake of conciseness, we have chosen not to detail these applications, but instead recommend the interested reader to the above articles and the references therein.

4

Fine-grained spectral analysis: ℓ_∞ and $\ell_{2,\infty}$ perturbation theory

In a growing number of applications, the ℓ_2 -type distance between subspaces, which is the central subject studied in Chapter 3, turns out to be inadequate for performance characterization. Rather, what would be of interest is the entrywise behavior of the eigenvector and the matrix under consideration. This is especially important when the individual entries of the eigenvector or the matrix of interest carry pivotal operational meanings. For example, in a recommendation system, one might be interested in controlling the prediction error of a user's preference on a specific product, which concerns a specific entry in a user-product rating matrix; in sensor network localization, one might seek to control the ranging error w.r.t. a pair of sensors, which corresponds to entrywise prediction errors in a Euclidean distance matrix; and last but not least, in community recovery, the entries of the leading eigenvector of a certain data matrix might encode the community membership associated with each individual (as explained in Section 3.4).

Tackling the preceding applications calls for development of fine-grained spectral analysis beyond classical ℓ_2 perturbation theory. To be more precise, consider once again the observation model

$$\mathbf{M} = \mathbf{M}^* + \mathbf{E}$$

previously studied in Chapter 2 (cf. (2.1)). The sort of fine-grained theory being sought after gravitates around the following questions concerned with ℓ_∞ and/or $\ell_{2,\infty}$ perturbation:

- For a symmetric matrix \mathbf{M}^* , how to characterize the effect of \mathbf{E} on the ℓ_∞ perturbation of the leading eigenvector, or the $\ell_{2,\infty}$ perturbation of the rank- r leading eigenspace?
- For a general matrix \mathbf{M}^* , how to pin down the ℓ_∞ perturbation of the leading singular vector, or the $\ell_{2,\infty}$ perturbation of the rank- r leading singular subspace, in response to the perturbation \mathbf{E} ?
- How to assess the entrywise estimation error of the matrix estimate produced by the spectral method, and how is it affected by \mathbf{E} ?

Unfortunately, a direct application of classical ℓ_2 perturbation theory typically leads to overly crude bounds when coping with the above questions. In particular, when the ℓ_2 error is approximately evenly distributed across entries, naively upper bounding the entrywise error by the ℓ_2 error is often loose by an order-of-magnitude. In order to conquer such limitations, this chapter introduces a modern suite of techniques that delivers tight ℓ_∞ and $\ell_{2,\infty}$ error control by leveraging the statistical nature of data models.

4.1 Leave-one-out analysis: An illustrative example

To paint a high-level picture of the core ideas empowering the ℓ_∞ and $\ell_{2,\infty}$ analysis, we find it helpful to first look at a pedagogical example of rank-1 matrix denoising, a special case of the formulation introduced in Section 3.2.1.

4.1.1 Setup and algorithm

Suppose that we observe a noisy copy of an unknown rank-1 matrix \mathbf{M}^* as follows

$$\mathbf{M} = \mathbf{M}^* + \mathbf{E} = \lambda^* \mathbf{u}^* \mathbf{u}^{*\top} + \mathbf{E} \in \mathbb{R}^{n \times n}, \quad (4.1)$$

where $\lambda^* > 0$ and $\mathbf{u}^* \in \mathbb{R}^n$ represent the largest eigenvalue of \mathbf{M}^* and its associated eigenvector, respectively. We assume the Gaussian

noise model as in Section 3.2.1, namely, \mathbf{E} is a symmetric matrix whose upper triangular part comprises of i.i.d. entries drawn from $\mathcal{N}(0, \sigma^2)$. In addition, we remind the readers of the following incoherence parameter μ (cf. Definition 3.1):

$$\mu = n\|\mathbf{u}^*\|_\infty^2, \quad (4.2)$$

which satisfies $1 \leq \mu \leq n$ in this rank-1 case; see Remark 3.12.

Letting λ be the leading eigenvalue of \mathbf{M} (i.e., $\lambda = \lambda_1(\mathbf{M})$) and $\mathbf{u} \in \mathbb{R}^n$ the associated eigenvector, the spectral method attempts to estimate \mathbf{u}^* using \mathbf{u} . In this section, we are particularly interested in controlling the entrywise error, defined in terms of the ℓ_∞ distance (modulo the global sign):

$$\text{dist}_\infty(\mathbf{u}, \mathbf{u}^*) := \min \{\|\mathbf{u} - \mathbf{u}^*\|_\infty, \|\mathbf{u} + \mathbf{u}^*\|_\infty\}. \quad (4.3)$$

4.1.2 ℓ_∞ performance guarantees

While Section 3.2.2 delivers ℓ_2 estimation guarantees for the spectral estimate \mathbf{u} , it falls short of characterizing the entrywise behavior—except for the crude and highly suboptimal bound $\text{dist}_\infty(\mathbf{u}, \mathbf{u}^*) \leq \text{dist}(\mathbf{u}, \mathbf{u}^*)$. Encouragingly, this simple spectral method is provably accurate in an entrywise fashion, as revealed by the following theorem.

Theorem 4.1. Consider the settings in Section 4.1.1. There exists some sufficiently small constant $c_0 > 0$ such that if $\sigma\sqrt{n} \leq c_0\lambda^*$, then

$$\text{dist}_\infty(\mathbf{u}, \mathbf{u}^*) \lesssim \frac{\sigma(\sqrt{\log n} + \sqrt{\mu})}{\lambda^*} \quad (4.4)$$

holds with probability exceeding $1 - O(n^{-8})$.

In particular, if the incoherence parameter obeys $\mu \lesssim \log n$ (the case where no entries of \mathbf{u}^* are significantly larger in magnitude than the average magnitude), then our ℓ_∞ bound reads

$$\text{dist}_\infty(\mathbf{u}, \mathbf{u}^*) \lesssim \frac{\sigma\sqrt{\log n}}{\lambda^*}, \quad (4.5)$$

which is about $\sqrt{n/\log n}$ times smaller than the ℓ_2 error bound (3.13), that is, $\text{dist}(\mathbf{u}, \mathbf{u}^*) \lesssim \frac{\sigma\sqrt{n}}{\lambda^*}$. This implies that the estimation errors of \mathbf{u} are dispersed more or less evenly across all entries—a message that is previously unavailable from classical ℓ_2 perturbation theory.

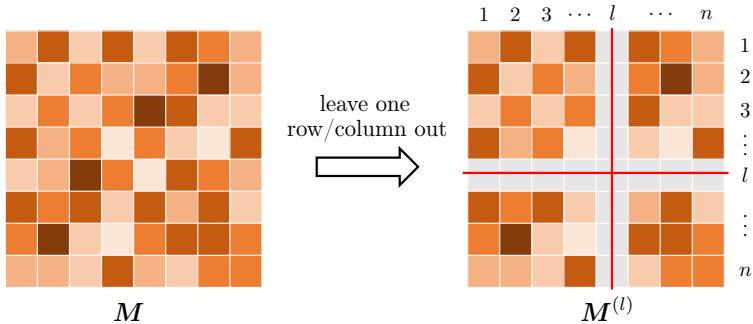


Figure 4.1: Illustration of the leave-one-out auxiliary matrix $\mathbf{M}^{(l)}$, which removes all the noise in the l -th row and the l -th column of \mathbf{M} .

4.1.3 Key ingredient and intuition: Leave-one-out estimates

To facilitate entrywise analysis, a crucial ingredient lies in the introduction of a set of leave-one-out auxiliary estimates, detailed below.

Construction of leave-one-out estimates. For each $1 \leq l \leq n$, let us construct an auxiliary matrix $\mathbf{M}^{(l)}$ as follows

$$\mathbf{M}^{(l)} := \lambda^* \mathbf{u}^* \mathbf{u}^{*\top} + \mathbf{E}^{(l)}, \quad (4.6)$$

where the noise matrix $\mathbf{E}^{(l)}$ is generated according to

$$E_{i,j}^{(l)} := \begin{cases} E_{i,j}, & \text{if } i \neq l \text{ and } j \neq l, \\ 0, & \text{else.} \end{cases} \quad (4.7)$$

In words, $\mathbf{M}^{(l)}$ (resp. $\mathbf{E}^{(l)}$) is obtained by leaving out the randomness in the l -th row/column of \mathbf{M} (resp. \mathbf{E}). The pattern of the leave-one-out construction is illustrated in Figure 4.1. Let $\lambda^{(l)}$ and $\mathbf{u}^{(l)}$ denote respectively the leading eigenvalue and leading eigenvector of $\mathbf{M}^{(l)}$; these leave-one-out estimates are introduced solely for analysis purpose. It is important to recognize that by construction, $\mathbf{M}^{(l)}$ (and hence $\mathbf{u}^{(l)}$) is independent of the noise in the l -th row/column of \mathbf{E} , a fact that plays a pivotal role in controlling the perturbation of the l -th entry of \mathbf{u}^* .

Intuition. Before delving into the proof, let us first explain the rationale at an intuitive level.

- Given that $\mathbf{u}^{(l)}$ is obtained by dropping only a tiny fraction of the data, we expect $\mathbf{u}^{(l)}$ to be exceedingly close to \mathbf{u} , i.e.,

$$\mathbf{u} \approx \pm \mathbf{u}^{(l)}. \quad (4.8)$$

In words, $\mathbf{u}^{(l)}$ forms a reliable surrogate of \mathbf{u} , which motivates us to analyze $\mathbf{u}^{(l)}$ instead (if there are foreseeable benefits to do so).

- The way we construct $\mathbf{u}^{(l)}$ makes it particularly convenient to analyze the behavior of the l -th entry, denoted by $u_l^{(l)}$. More specifically, given that $(\lambda^{(l)}, \mathbf{u}^{(l)})$ is an eigenpair of $\mathbf{M}^{(l)}$, one has (assuming for the moment that $\lambda^{(l)} \neq 0$)

$$\begin{aligned} u_l^{(l)} &= \frac{1}{\lambda^{(l)}} \mathbf{M}_{l,\cdot}^{(l)} \mathbf{u}^{(l)} = \frac{1}{\lambda^{(l)}} \mathbf{M}_{l,\cdot}^* \mathbf{u}^{(l)} = \frac{\lambda^*}{\lambda^{(l)}} u_l^* \mathbf{u}^{*\top} \mathbf{u}^{(l)} \\ &\approx \pm u_l^*. \end{aligned} \quad (4.9)$$

Here, the first line follows since, by design, the l -th rows of $\mathbf{M}^{(l)}$ and \mathbf{M}^* coincide (both of which are given by $\lambda^* u_l^* \mathbf{u}^{*\top}$), whereas the second line holds as long as the size σ of the noise is sufficiently small, so that $\lambda^{(l)}/\lambda^* \approx 1$ and $\mathbf{u}^{*\top} \mathbf{u}^{(l)} \approx \pm 1$ according to the ℓ_2 perturbation theory.

Combining the above observations suggests that $u_l \approx \pm u_l^{(l)} \approx \pm u_l^*$.

4.1.4 Leave-one-out analysis

Now we make rigorous the heuristic argument in the last subsection, which relies heavily on careful statistical analysis.

Preparation: what we have learned from ℓ_2 perturbation theory

Let us start by collecting a few results derived from the ℓ_2 perturbation theory in Section 3.2 for handy reference. Experienced readers can proceed directly to Step 1.

Specifically, suppose that $\sigma\sqrt{n} \leq \frac{1-1/\sqrt{2}}{5} \lambda^*$. Then with probability at least $1 - O(n^{-8})$,

$$\|\mathbf{E}\| \leq 5\sigma\sqrt{n} \quad \|\mathbf{E}^{(l)}\| \leq \|\mathbf{E}\| \leq 5\sigma\sqrt{n} \quad (4.11a)$$

$$\text{dist}(\mathbf{u}, \mathbf{u}^*) \leq \frac{10\sigma\sqrt{n}}{\lambda^*} \quad \text{dist}(\mathbf{u}^{(l)}, \mathbf{u}^*) \leq \frac{10\sigma\sqrt{n}}{\lambda^*} \quad (4.11b)$$

$$|\lambda - \lambda^*| \leq 5\sigma\sqrt{n} \quad |\lambda^{(l)} - \lambda^*| \leq 5\sigma\sqrt{n} \quad (4.11c)$$

$$\max_{j:j \geq 2} |\lambda_j(\mathbf{M})| \leq 5\sigma\sqrt{n} \quad \max_{j:j \geq 2} |\lambda_j(\mathbf{M}^{(l)})| \leq 5\sigma\sqrt{n} \quad (4.11d)$$

hold simultaneously for all $1 \leq l \leq n$. Here, the first line arises from (3.12), and the remaining claims follow the same argument as in the proof of Corollary 2.8. Consequently, there exist global signs $z, z_l \in \{1, -1\}$ obeying $\|z\mathbf{u} - \mathbf{u}^*\|_2 = \text{dist}(\mathbf{u}, \mathbf{u}^*) \leq 10\sigma\sqrt{n}/\lambda^*$ and $\|z_l\mathbf{u}^{(l)} - \mathbf{u}^*\|_2 = \text{dist}(\mathbf{u}^{(l)}, \mathbf{u}^*) \leq 10\sigma\sqrt{n}/\lambda^*$. To simplify presentation, we shall assume

$$\|\mathbf{u} - \mathbf{u}^*\|_2 = \text{dist}(\mathbf{u}, \mathbf{u}^*), \quad (4.12a)$$

$$\|\mathbf{u}^{(l)} - \mathbf{u}^*\|_2 = \text{dist}(\mathbf{u}^{(l)}, \mathbf{u}^*), \quad 1 \leq l \leq n \quad (4.12b)$$

without loss of generality. As a simple yet useful byproduct: if $20\sigma\sqrt{n} < \lambda^*$, then Condition (4.12) necessarily implies

$$\|\mathbf{u} - \mathbf{u}^{(l)}\|_2 = \text{dist}(\mathbf{u}, \mathbf{u}^{(l)}), \quad 1 \leq l \leq n. \quad (4.13)$$

To see this, combine the triangle inequality and (4.11) to yield

$$\|\mathbf{u} - \mathbf{u}^{(l)}\|_2 \leq \|\mathbf{u} - \mathbf{u}^*\|_2 + \|\mathbf{u}^{(l)} - \mathbf{u}^*\|_2 \leq \frac{20\sigma\sqrt{n}}{\lambda^*} < 1,$$

which taken collectively with the fact $\|\mathbf{u}\|_2 = \|\mathbf{u}^{(l)}\|_2 = 1$ gives

$$\|\mathbf{u} + \mathbf{u}^{(l)}\|_2^2 = 2\|\mathbf{u}\|_2^2 + 2\|\mathbf{u}^{(l)}\|_2^2 - \|\mathbf{u} - \mathbf{u}^{(l)}\|_2^2 > 1 > \|\mathbf{u} - \mathbf{u}^{(l)}\|_2^2.$$

This together with the definition (2.10a) of $\text{dist}(\cdot, \cdot)$ validates (4.13).

Step 1: bounding the proximity of leave-one-out & true estimates

In this step, we seek to control the distance between the true estimate \mathbf{u} and the leave-one-out estimate $\mathbf{u}^{(l)}$. Suppose for the moment that

$$\|\mathbf{M} - \mathbf{M}^{(l)}\| \leq (1 - 1/\sqrt{2}) \left(\lambda^{(l)} - \max_{j \geq 2} |\lambda_j(\mathbf{M}^{(l)})| \right). \quad (4.14)$$

We can then invoke the Davis-Kahan theorem (cf. Corollary 2.8) and the relation (4.13) to yield

$$\|\mathbf{u} - \mathbf{u}^{(l)}\|_2 \leq \frac{2\|(\mathbf{M} - \mathbf{M}^{(l)})\mathbf{u}^{(l)}\|_2}{\lambda^{(l)} - \max_{j \geq 2} |\lambda_j(\mathbf{M}^{(l)})|} \leq \frac{4\|(\mathbf{M} - \mathbf{M}^{(l)})\mathbf{u}^{(l)}\|_2}{\lambda^*}. \quad (4.15)$$

Here, the last inequality invokes (4.11) and $20\sigma\sqrt{n} \leq \lambda^*$ to obtain

$$\lambda^{(l)} - \max_{j \geq 2} |\lambda_j(\mathbf{M}^{(l)})| \geq (\lambda^* - 5\sigma\sqrt{n}) - 5\sigma\sqrt{n} \geq \lambda^*/2. \quad (4.16)$$

A byproduct of this calculation is that $\lambda^{(l)}$ is positive for all $1 \leq l \leq n$.

It thus remains to control the term $\|(\mathbf{M} - \mathbf{M}^{(l)})\mathbf{u}^{(l)}\|_2$ in (4.15), towards which certain statistical independence proves crucial. Specifically, we observe that (by construction of $\mathbf{M}^{(l)}$)

$$(\mathbf{M} - \mathbf{M}^{(l)})\mathbf{u}^{(l)} = \mathbf{e}_l \mathbf{E}_{l,\cdot} \mathbf{u}^{(l)} + u_l^{(l)} (\mathbf{E}_{\cdot,l} - E_{l,l} \mathbf{e}_l),$$

where \mathbf{e}_l is the l -th standard basis vector, and $\mathbf{E}_{l,\cdot}$ (resp. $\mathbf{E}_{\cdot,l}$) denotes the l -th row (resp. column) of \mathbf{E} . By construction, $\mathbf{u}^{(l)}$ is statistically independent of $\mathbf{E}_{l,\cdot}$, thus indicating that

$$\mathbf{E}_{l,\cdot} \mathbf{u}^{(l)} \sim \mathcal{N}(0, \sigma^2 \|\mathbf{u}^{(l)}\|_2^2) = \mathcal{N}(0, \sigma^2) \quad (4.17)$$

conditioned on $\mathbf{u}^{(l)}$. Hence, with probability at least $1 - n^{-10}$,

$$|\mathbf{E}_{l,\cdot} \mathbf{u}^{(l)}| \leq 5\sigma\sqrt{\log n}, \quad 1 \leq l \leq n. \quad (4.18)$$

In addition, $\|\mathbf{E}_{\cdot,l} - E_{l,l} \mathbf{e}_l\|_2 \leq \|\mathbf{E}_{\cdot,l}\|_2 \leq \|\mathbf{E}\| \leq 5\sigma\sqrt{n}$ (cf. (4.11)). Consequently,

$$\begin{aligned} \|(\mathbf{M} - \mathbf{M}^{(l)})\mathbf{u}^{(l)}\|_2 &\leq |\mathbf{E}_{l,\cdot} \mathbf{u}^{(l)}| + \|\mathbf{E}_{\cdot,l}\|_2 \cdot |u_l^{(l)}| \\ &\leq 5\sigma\sqrt{\log n} + \|\mathbf{E}_{\cdot,l}\|_2 (|u_l| + \|\mathbf{u} - \mathbf{u}^{(l)}\|_\infty) \\ &\leq 5\sigma\sqrt{\log n} + 5\sigma\sqrt{n} \|\mathbf{u}\|_\infty + 5\sigma\sqrt{n} \|\mathbf{u} - \mathbf{u}^{(l)}\|_2. \end{aligned}$$

Substitution into (4.15) yields

$$\begin{aligned} \|\mathbf{u} - \mathbf{u}^{(l)}\|_2 &\leq \frac{20\sigma\sqrt{\log n} + 20\sigma\sqrt{n} \|\mathbf{u}\|_\infty + 20\sigma\sqrt{n} \|\mathbf{u} - \mathbf{u}^{(l)}\|_2}{\lambda^*} \\ &\leq \frac{20\sigma\sqrt{\log n} + 20\sigma\sqrt{n} \|\mathbf{u}\|_\infty}{\lambda^*} + \frac{1}{2} \|\mathbf{u} - \mathbf{u}^{(l)}\|_2, \end{aligned}$$

provided that $40\sigma\sqrt{n} \leq \lambda^*$. Rearranging terms and taking the union bound, we demonstrate that with probability at least $1 - O(n^{-8})$,

$$\|\mathbf{u} - \mathbf{u}^{(l)}\|_2 \leq \frac{40\sigma\sqrt{\log n} + 40\sigma\sqrt{n} \|\mathbf{u}\|_\infty}{\lambda^*}, \quad 1 \leq l \leq n. \quad (4.19)$$

Proof of the relation (4.14). Apply the triangle inequality to see that

$$\begin{aligned}\|\mathbf{M} - \mathbf{M}^{(l)}\| &\leq \|\mathbf{M} - \mathbf{M}^*\| + \|\mathbf{M}^* - \mathbf{M}^{(l)}\| = \|\mathbf{E}\| + \|\mathbf{E}^{(l)}\| \\ &\leq 10\sigma\sqrt{n} \leq \lambda^*/10.\end{aligned}$$

Here, the equality arises from the definition of \mathbf{M} and \mathbf{M}^* , the penultimate inequality uses (4.11), while the last inequality holds as long as $100\sigma\sqrt{n} \leq \lambda^*$. This together with (4.16) establishes (4.14). \square

Step 2: analyzing leave-one-out estimates

We now turn attention to bounding the size of the l -th entry $u_l^{(l)}$ of $\mathbf{u}^{(l)}$. Since $\lambda^{(l)} > 0$, by (4.9), we have

$$\begin{aligned}u_l^{(l)} - u_l^* &= u_l^* \left(\frac{\lambda^*}{\lambda^{(l)}} \mathbf{u}^{*\top} \mathbf{u}^{(l)} - \mathbf{u}^{*\top} \mathbf{u}^* \right) \\ &= u_l^* \left(\frac{\lambda^* - \lambda^{(l)}}{\lambda^{(l)}} \mathbf{u}^{*\top} \mathbf{u}^{(l)} \right) + u_l^* \mathbf{u}^{*\top} (\mathbf{u}^{(l)} - \mathbf{u}^*).\end{aligned}$$

The triangle inequality and the Cauchy-Schwarz inequality then give

$$\begin{aligned}|u_l^{(l)} - u_l^*| &\leq |u_l^*| \cdot \frac{|\lambda^* - \lambda^{(l)}|}{|\lambda^{(l)}|} \cdot \|\mathbf{u}^*\|_2 \cdot \|\mathbf{u}^{(l)}\|_2 + |u_l^*| \cdot \|\mathbf{u}^*\|_2 \cdot \|\mathbf{u}^{(l)} - \mathbf{u}^*\|_2 \\ &\leq |u_l^*| \cdot \frac{10\sigma\sqrt{n}}{\lambda^*} + |u_l^*| \cdot \frac{10\sigma\sqrt{n}}{\lambda^*} \\ &\leq \frac{20\sigma\sqrt{n}}{\lambda^*} \|\mathbf{u}^*\|_\infty.\end{aligned}\tag{4.20}$$

Here, the second line holds due to Condition (4.11) and the fact $|\lambda^{(l)}| \geq \lambda^*/2$; see (4.16).

Step 3: putting all pieces together

Putting (4.19) and (4.20) together, we arrive at

$$\begin{aligned}\|\mathbf{u} - \mathbf{u}^*\|_\infty &= \max_l |u_l - u_l^*| \leq \max_l \left\{ |u_l^{(l)} - u_l^*| + \|\mathbf{u} - \mathbf{u}^{(l)}\|_2 \right\} \\ &\leq \frac{20\sigma\sqrt{n}}{\lambda^*} \|\mathbf{u}^*\|_\infty + \frac{40\sigma\sqrt{\log n} + 40\sigma\sqrt{n} \|\mathbf{u}\|_\infty}{\lambda^*}.\end{aligned}\tag{4.21}$$

The above upper bound, however, involves the term $\|\mathbf{u}\|_\infty$, which can further be bounded by $\|\mathbf{u}^*\|_\infty + \|\mathbf{u} - \mathbf{u}^*\|_\infty$. Substituting this into (4.21), we have

$$\|\mathbf{u} - \mathbf{u}^*\|_\infty \leq \frac{40\sigma\sqrt{\log n} + 60\sigma\sqrt{n}\|\mathbf{u}^*\|_\infty}{\lambda^*} + \frac{1}{2}\|\mathbf{u} - \mathbf{u}^*\|_\infty,$$

provided that $80\sigma\sqrt{n} \leq \lambda^*$. Rearranging terms yields

$$\|\mathbf{u} - \mathbf{u}^*\|_\infty \leq \frac{80\sigma\sqrt{\log n} + 120\sigma\sqrt{n}\|\mathbf{u}^*\|_\infty}{\lambda^*} = \frac{80\sigma\sqrt{\log n} + 120\sigma\sqrt{\mu}}{\lambda^*},$$

where the last identity results from the definition (4.2) of μ .

4.2 $\ell_{2,\infty}$ eigenspace perturbation under independent noise

The appealing entrywise behavior of the eigenvector estimator in Section 4.1 hints at the promising performance of spectral methods for broader contexts. In this section, we set out to develop a more general framework about $\ell_{2,\infty}$ eigenspace perturbation that covers a wide spectrum of scenarios.

4.2.1 Setup and notation

Ground truth. Consider a rank- r symmetric matrix $\mathbf{M}^* \in \mathbb{R}^{n \times n}$ with eigenvectors $\mathbf{u}_1^*, \dots, \mathbf{u}_n^*$ and associated eigenvalues $\lambda_1^*, \dots, \lambda_n^*$ obeying

$$|\lambda_1^*| \geq |\lambda_2^*| \geq \dots \geq |\lambda_r^*| > 0 \quad \text{and} \quad \lambda_{r+1}^* = \dots = \lambda_n^* = 0. \quad (4.22)$$

We shall write the eigendecomposition $\mathbf{M}^* = \mathbf{U}^* \mathbf{\Lambda}^* \mathbf{U}^{*\top}$ as usual, where $\mathbf{\Lambda}^* := \text{diag}([\lambda_1^*, \dots, \lambda_r^*])$ and $\mathbf{U}^* := [\mathbf{u}_1^*, \dots, \mathbf{u}_r^*] \in \mathbb{R}^{n \times r}$. Denote the condition number of \mathbf{M}^* as

$$\kappa := |\lambda_1^*| / |\lambda_r^*|. \quad (4.23)$$

Akin to Definition 3.1, the incoherence parameter of \mathbf{M}^* is defined as

$$\mu := \frac{n\|\mathbf{U}^*\|_{2,\infty}^2}{r}, \quad (4.24)$$

a parameter that captures how well the energy of \mathbf{U}^* is spread out across all rows and that obeys (see Remark 3.12)

$$1 \leq \mu \leq n/r. \quad (4.25)$$

Observed data. What we observe is a corrupted version

$$\mathbf{M} = \mathbf{M}^* + \mathbf{E} \in \mathbb{R}^{n \times n}, \quad (4.26)$$

where \mathbf{E} is a symmetric noise matrix. We denote by $\{\lambda_i\}_{1 \leq i \leq n}$ the set of eigenvalues of \mathbf{M} obeying

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|, \quad (4.27)$$

and let \mathbf{u}_i be the eigenvector of \mathbf{M} associated with λ_i . We shall introduce the diagonal matrix $\mathbf{\Lambda} \in \mathbb{R}^{r \times r}$ as $\mathbf{\Lambda} := \text{diag}([\lambda_1, \dots, \lambda_r])$.

Noise assumptions. This section aims to cover a fairly broad class of scenarios of independent noise. In particular, the noise matrix considered herein is assumed to satisfy the mild conditions listed below.

Assumption 4.1. The entries in the lower triangular part of $\mathbf{E} = [E_{i,j}]_{1 \leq i,j \leq n}$ are independently generated obeying

$$\mathbb{E}[E_{i,j}] = 0, \quad \mathbb{E}[E_{i,j}^2] =: \sigma_{i,j}^2 \leq \sigma^2, \quad |E_{i,j}| \leq B, \quad \text{for all } i \geq j. \quad (4.28)$$

In particular, σ^2 is taken to be the smallest choice satisfying (4.28). Further, it is assumed that

$$c_b := \frac{B}{\sigma \sqrt{n/(\mu \log n)}} = O(1). \quad (4.29)$$

We emphasize that both σ and B are quantities that are allowed to scale with n . When μ is not too large, Condition (4.29) allows the maximum magnitude B of each noisy entry to be substantially larger than the typical size σ .

Goal and algorithm. We seek to estimate \mathbf{U}^* based on \mathbf{M} . Towards this, a simple spectral method computes the matrix $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_r] \in \mathbb{R}^{n \times r}$ that comprises the top- r leading eigenvectors of \mathbf{M} .

4.2.2 $\ell_{2,\infty}$ and ℓ_∞ theoretical guarantees

The leave-one-out argument introduced before, when properly strengthened, enables powerful $\ell_{2,\infty}$ performance guarantees for the spectral

estimate \mathbf{U} , which concern row-wise perturbation of the eigenspace. Before continuing, we remind the readers of the global rotation ambiguity, namely, in general we cannot expect \mathbf{U} to be close to \mathbf{U}^* unless suitable global rotation is taken into account. In light of this, we introduce the following notation that helps identify a proper rotation matrix.

Definition 4.1. For any matrix \mathbf{Z} with SVD $\mathbf{Z} = \mathbf{U}_Z \boldsymbol{\Sigma}_Z \mathbf{V}_Z^\top$ (where \mathbf{U}_Z and \mathbf{V}_Z represent respectively the left and right singular matrices of \mathbf{Z} , and $\boldsymbol{\Sigma}_Z$ is a diagonal matrix composed of the singular values), define

$$\text{sgn}(\mathbf{Z}) := \mathbf{U}_Z \mathbf{V}_Z^\top \quad (4.30)$$

to be the matrix sign function of \mathbf{Z} .

Remark 4.1. The matrix sign function is commonly encountered when aligning two matrices—classically known as the orthogonal Procrustes problem (Schönenmann, 1966). Consider any two matrices $\widehat{\mathbf{B}}, \mathbf{B} \in \mathbb{R}^{n \times r}$ with $r \leq n$. Among all rotation matrices, the one that best aligns $\widehat{\mathbf{B}}$ with \mathbf{B} is precisely $\text{sgn}(\widehat{\mathbf{B}}^\top \mathbf{B})$ (see, e.g., (Ma *et al.*, 2020, Appendix D.2.1)), namely,

$$\text{sgn}(\widehat{\mathbf{B}}^\top \mathbf{B}) = \arg \min_{\mathbf{O} \in \mathcal{O}^{r \times r}} \|\widehat{\mathbf{B}}\mathbf{O} - \mathbf{B}\|_F^2.$$

With this definition in place, we are ready to state an $\ell_{2,\infty}$ theory adapted from Abbe *et al.* (2020b). Compared to the original development in Abbe *et al.* (2020b), the theorem and its proof provided herein are more streamlined versions tailored to the current setting.

Theorem 4.2. Consider the settings and assumptions in Section 4.2.1. Define $\mathbf{H} := \mathbf{U}^\top \mathbf{U}^*$. With probability exceeding $1 - O(n^{-5})$, one has

$$\|\mathbf{U} \text{sgn}(\mathbf{H}) - \mathbf{U}^*\|_{2,\infty} \lesssim \frac{\sigma \kappa \sqrt{\mu r} + \sigma \sqrt{r \log n}}{|\lambda_r^*|}, \quad (4.31a)$$

$$\begin{aligned} \|\mathbf{U} \text{sgn}(\mathbf{H}) - \mathbf{M} \mathbf{U}^* (\boldsymbol{\Lambda}^*)^{-1}\|_{2,\infty} \\ \lesssim \frac{\sigma \kappa \sqrt{\mu r}}{|\lambda_r^*|} + \frac{\sigma^2 \sqrt{rn \log n} + \sigma B \sqrt{\mu r \log^3 n}}{(\lambda_r^*)^2}, \end{aligned} \quad (4.31b)$$

provided that $\sigma \sqrt{n \log n} \leq c_\sigma |\lambda_r^*|$ for some sufficiently small constant $c_\sigma > 0$.

The proof of this theorem can be found in Section 4.8. Note that under the assumption in the theorem, the bound on the right-hand side of (4.31b) is no larger than the one on the right-hand side of (4.31a). In fact, (4.31b) could indeed be tighter than (4.31a) in some important scenarios like community recovery (see Section 4.5).

The $\ell_{2,\infty}$ perturbation theory in Theorem 4.2 accommodates a broad family of noise matrices with independent entries. In the sequel, we take a moment to interpret several key messages conveyed by this result.

De-localization of estimation errors. For simplicity, let us concentrate on the case where $\mu, \kappa = O(1)$. Note that the Davis-Kahan theorem introduced previously results in the following ℓ_2 estimation guarantees (to be detailed in Section 4.8.2)

$$\text{dist}_F(\mathbf{U}, \mathbf{U}^*) \leq \sqrt{r} \text{dist}(\mathbf{U}, \mathbf{U}^*) \lesssim \frac{\sigma \sqrt{nr}}{|\lambda_r^*|}. \quad (4.32)$$

In comparison, the $\ell_{2,\infty}$ bound derived in Theorem 4.2 simplifies to

$$\min_{\mathbf{R} \in \mathcal{O}^{r \times r}} \|\mathbf{U}\mathbf{R} - \mathbf{U}^*\|_{2,\infty} \leq \|\mathbf{U}\text{sgn}(\mathbf{H}) - \mathbf{U}^*\|_{2,\infty} \lesssim \frac{\sigma \sqrt{r \log n}}{|\lambda_r^*|} \quad (4.33)$$

under the condition $\mu, \kappa = O(1)$, which is about $O(\sqrt{n/\log n})$ times smaller than the Euclidean error bound (4.32). This implies that the estimation error of \mathbf{U} is fairly de-localized and spread out across all rows.

First-order approximation. Informally, Theorem 4.2 (and its analysis) unveils the goodness of the first-order approximation

$$\mathbf{U}\text{sgn}(\mathbf{H}) \approx \mathbf{M}\mathbf{U}^*(\Lambda^*)^{-1} = \mathbf{U}^* + \mathbf{E}\mathbf{U}^*(\Lambda^*)^{-1} \quad (4.34)$$

uniformly across all rows. An implication of Theorem 4.2 is that \mathbf{U} might be closer to the first-order approximation $\mathbf{M}\mathbf{U}^*(\Lambda^*)^{-1}$ than to the ground truth \mathbf{U}^* (namely, the upper bound on the right-hand side of (4.31b) is smaller than the bound on the right-hand side of (4.31a) under the stated conditions). In principle, the linear term $\mathbf{E}\mathbf{U}^*(\Lambda^*)^{-1}$ can be viewed as a correction term that helps improve the approximation fidelity. As we shall see momentarily in Section 4.5, this subtle difference

leads to sharper performance guarantees in applications like community recovery.

Entrywise estimation errors. There is no shortage of applications where one cares more about the fine-grained estimation accuracy of the matrix rather than that of the low-rank factors. Fortunately, the $\ell_{2,\infty}$ theory derived in Theorem 4.2 in turn enables entrywise performance guarantees when estimating the matrix \mathbf{M}^* . This is summarized in the following corollary, with the proof deferred to Section 4.9.

Corollary 4.3. Consider the settings and assumptions in Section 4.2.1, and assume further that $\sigma\kappa\sqrt{n\log n} \leq c_1|\lambda_r^*|$ for some sufficiently small constant $c_1 > 0$. Then with probability at least $1 - O(n^{-5})$, one has

$$\|\mathbf{U}\Lambda\mathbf{U}^\top - \mathbf{M}^*\|_\infty \lesssim \sigma\kappa^2\mu r\sqrt{\frac{\log n}{n}}. \quad (4.35)$$

Once again, it is instrumental to explain the result by specializing it to the simpler regime where $\kappa, \mu, r = O(1)$. In this case, the finding of Corollary 4.3 reduces to

$$\|\mathbf{U}\Lambda\mathbf{U}^\top - \mathbf{M}^*\|_\infty \lesssim \sigma\sqrt{\frac{\log n}{n}}. \quad (4.36)$$

In comparison, the Euclidean error of this spectral estimate satisfies (which follows by combining (3.15) with (3.9) and (4.29))

$$\|\mathbf{U}\Lambda\mathbf{U}^\top - \mathbf{M}^*\|_{\text{F}} \leq 2\sqrt{2}\|\mathbf{E}\| \lesssim \sigma\sqrt{n} \quad (4.37)$$

with high probability, which is on the order of $n/\sqrt{\log n}$ times larger than the entrywise error bound (4.36). In other words, the energy of the estimation error of the unknown matrix is also dispersed more or less across all matrix entries, a message that cannot be derived from classical matrix perturbation theory alone.

Leave-one-out analysis. As alluded to previously, the proof of Theorem 4.2 relies heavily upon the leave-one-out analysis framework to decouple delicate statistical dependency. While the core idea bears close resemblance to the exposition in Section 4.1.4, implementing this idea rigorously for the general case requires considerably more effort. We defer a complete proof to Section 4.8.

4.3 $\ell_{2,\infty}$ singular subspace perturbation under independent noise

The general theory presented in Section 4.2 applies only to symmetric matrices. It is not uncommon, however, to encounter scenarios where the matrix of interest \mathbf{M}^* is asymmetric. This motivates the need of extending the $\ell_{2,\infty}$ perturbation theory to accommodate more general matrices, which is the main content of the current section.

4.3.1 Setup and notation

Ground truth. Consider an unknown rank- r matrix $\mathbf{M}^* \in \mathbb{R}^{n_1 \times n_2}$. Let $\mathbf{M}^* = \mathbf{U}^* \mathbf{\Sigma}^* \mathbf{V}^{*\top}$ represent the SVD of \mathbf{M}^* , where $\mathbf{U}^* \in \mathbb{R}^{n_1 \times r}$ (resp. $\mathbf{V}^* \in \mathbb{R}^{n_2 \times r}$) entails the top- r left (resp. right) singular vectors of \mathbf{M}^* , and $\mathbf{\Sigma}^* = \text{diag}([\sigma_1^*, \sigma_2^*, \dots, \sigma_r^*])$ is formed by the (nonzero) singular values of \mathbf{M}^* . We arrange the singular values $\{\sigma_i^*\}$ in descending order (i.e., $\sigma_1^* \geq \sigma_2^* \geq \dots \geq \sigma_r^* > 0$).

Key parameters. As usual, μ stands for the incoherence parameter of \mathbf{M}^* (see Definition 3.1), and the condition number of the matrix \mathbf{M}^* is defined as $\kappa := \sigma_1^*/\sigma_r^*$. Without loss of generality, it is assumed that

$$n_1 \leq n_2$$

and we set $n := n_1 + n_2$.

Observations and noise assumptions. Assume we have access to corrupted observations of \mathbf{M}^* as follows:

$$\mathbf{M} = \mathbf{M}^* + \mathbf{E} \in \mathbb{R}^{n_1 \times n_2},$$

where $\mathbf{E} = [E_{i,j}]$ stands for a noise or perturbation matrix. We impose the following conditions on \mathbf{E} , which is a natural adaptation of Assumption 4.1 to the general case and covers a diverse array of scenarios.

Assumption 4.2. The entries of \mathbf{E} are independently generated obeying

$$\mathbb{E}[E_{i,j}] = 0, \quad \mathbb{E}[E_{i,j}^2] \leq \sigma^2, \quad |E_{i,j}| \leq B \quad \text{for all } i, j. \quad (4.38)$$

Further, assume that

$$c_b := \frac{B}{\sigma \sqrt{n_1/(\mu \log n)}} = O(1). \quad (4.39)$$

Goal and algorithm. Again, we aim at estimating \mathbf{U}^* and \mathbf{V}^* , based on the observation \mathbf{M} , using a spectral method. Specifically, let

$$\mathbf{M} = \begin{bmatrix} \mathbf{U} & \mathbf{U}_\perp \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma} & \\ & \boldsymbol{\Sigma}_\perp \end{bmatrix} \begin{bmatrix} \mathbf{U}^\top \\ \mathbf{V}^\top \end{bmatrix} \quad (4.40)$$

be the SVD of \mathbf{M} , in which $\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top$ is the rank- r SVD (i.e., the singular values in $\boldsymbol{\Sigma} := \text{diag}([\sigma_1, \dots, \sigma_r])$ are larger than those in $\boldsymbol{\Sigma}_\perp$). The spectral method then deploys (\mathbf{U}, \mathbf{V}) as an estimate of $(\mathbf{U}^*, \mathbf{V}^*)$.

4.3.2 $\ell_{2,\infty}$ and ℓ_∞ theoretical guarantees

We now present a theorem that generalizes Theorem 4.2 and Corollary 4.3 to accommodate general (asymmetric and possibly rectangular) matrices. This can be accomplished via a standard ‘‘symmetric dilation’’ trick; the details can be found in Section 4.10.

Theorem 4.4. Consider the settings and assumptions in Section 4.3.1, and define $\mathbf{H}_U := \mathbf{U}^\top \mathbf{U}^*$ and $\mathbf{H}_V := \mathbf{V}^\top \mathbf{V}^*$. With probability at least $1 - O(n^{-5})$, one has

$$\begin{aligned} \max \left\{ \|\mathbf{U} \text{sgn}(\mathbf{H}_U) - \mathbf{U}^*\|_{2,\infty}, \|\mathbf{V} \text{sgn}(\mathbf{H}_V) - \mathbf{V}^*\|_{2,\infty} \right\} \\ \lesssim \frac{\sigma \sqrt{r} (\kappa \sqrt{\frac{n_2}{n_1} \mu} + \sqrt{\log n})}{\sigma_r^*}, \end{aligned} \quad (4.41)$$

provided that $\sigma \sqrt{n \log n} \leq c_1 \sigma_r^*$ for some sufficiently small constant $c_1 > 0$. In addition, if $\sigma \kappa \sqrt{n \log n} \leq c_2 \sigma_r^*$ for some small enough constant $c_2 > 0$, then the following holds with probability at least $1 - O(n^{-5})$:

$$\|\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top - \mathbf{M}^*\|_\infty \lesssim \sigma \kappa^2 \mu r \sqrt{\frac{(n_2/n_1) \log n}{n_1}}. \quad (4.42)$$

The messages conveyed in Theorem 4.4 largely parallel those in Theorem 4.2 and Corollary 4.3. For simplicity, let us discuss the implications when $\kappa, \mu, r = O(1)$ and $n_1 \asymp n_2$ (i.e., the aspect ratio of the matrix is $n_2/n_1 = O(1)$). In this scenario, Theorem 4.4 implies that

$$\|\mathbf{U} \text{sgn}(\mathbf{H}_U) - \mathbf{U}^*\|_{2,\infty} + \|\mathbf{V} \text{sgn}(\mathbf{H}_V) - \mathbf{V}^*\|_{2,\infty} \lesssim \frac{\sigma \sqrt{\log n}}{\sigma_r^*},$$

$$\|\mathbf{U}\Sigma\mathbf{V}^\top - \mathbf{M}^*\|_\infty \lesssim \sigma \sqrt{\frac{\log n}{n}},$$

both of which bear close similarities to our previous observations (4.33) and (4.36). Akin to our discussions in Section 4.2.2, these findings tell us that the singular subspace estimation errors (resp. the matrix estimation errors) are fairly spread out across all rows of the singular subspace (resp. all entries of the matrix).

4.4 Application: Entrywise guarantees for matrix completion

To illustrate the utility of the fine-grained perturbation theory presented in previous sections, let us revisit the problem of matrix completion introduced in Section 3.8 and apply our refined theory.

As a recap, the spectral method proposed for matrix completion proceeds by computing the best rank- r approximation $\mathbf{U}\Sigma\mathbf{V}^\top$ of the rescaled data matrix $\mathbf{M} = p^{-1}\mathcal{P}_\Omega(\mathbf{M}^*)$, where p is the probability of each entry being observed, and $\mathcal{P}_\Omega(\cdot)$ denotes the Euclidean projection onto the set of matrices supported on the sampling set Ω . This time, we seek to characterize the $\ell_{2,\infty}$ error when estimating the true singular subspaces \mathbf{U}^* and \mathbf{V}^* , as well as the ℓ_∞ error when estimating the unknown matrix \mathbf{M}^* , as stated below. As before, we set $n := n_1 + n_2$.

Theorem 4.5. Consider the settings and assumptions in Section 3.8.1, and define $\mathbf{H}_U := \mathbf{U}^\top \mathbf{U}^*$ and $\mathbf{H}_V := \mathbf{V}^\top \mathbf{V}^*$. Suppose that $n_1 \leq n_2$ and $n_1 p \geq C\kappa^4 \mu^2 r^2 \log n$ for some sufficiently large constant $C > 0$. Then with probability greater than $1 - O(n^{-5})$, we have

$$\begin{aligned} \max\{\|\mathbf{U}\text{sgn}(\mathbf{H}_U) - \mathbf{U}^*\|_{2,\infty}, \|\mathbf{V}\text{sgn}(\mathbf{H}_V) - \mathbf{V}^*\|_{2,\infty}\} \\ \leq \kappa^2 \sqrt{\frac{\mu^3 r^3 \log n}{n_1^2 p}}; \end{aligned} \quad (4.43a)$$

$$\|\mathbf{U}\Sigma\mathbf{V}^\top - \mathbf{M}^*\|_\infty \lesssim \kappa^2 \mu^2 r^2 \sqrt{\frac{\log n}{n_1^3 p}} \|\mathbf{M}^*\|. \quad (4.43b)$$

Proof. Recall our notation $\mathbf{E} = \mathbf{M} - \mathbf{M}^* = p^{-1}\mathcal{P}_\Omega(\mathbf{M}^*) - \mathbf{M}^*$. It is straightforward to check that \mathbf{E} satisfies Assumption 4.2 with

$$\sigma^2 := \frac{\|\mathbf{M}^*\|_\infty^2}{p}, \quad \text{and} \quad B := \frac{\|\mathbf{M}^*\|_\infty}{p}. \quad (4.44)$$

In addition, from the relation $B = c_b \sigma \sqrt{n_1 / (\mu \log n)}$, it is seen that $c_b = O(1)$ holds as long as $n_1 p \gtrsim \mu \log n$. With these preparations in place, the claims in Theorem 4.5 follow directly from Theorem 4.4 and the bound (3.105b) on $\|\mathbf{M}^*\|_\infty$ (and hence on σ). \square

In what follows, we compare the $\ell_{2,\infty}$ and ℓ_2 performance guarantees derived in the above theorem with those ℓ_2 guarantees presented in Section 3.8.3; see Figure 1.4 in Section 1.1 for empirical performances. For the sake of brevity, we shall concentrate on the case where $\mu, \kappa, r = O(1)$.

- *$\ell_{2,\infty}$ singular space perturbation bounds.* Comparing the $\ell_{2,\infty}$ performance guarantee (4.43a) with Theorem 3.22, one sees that the $\ell_{2,\infty}$ perturbation bounds could be an order of $\sqrt{n_1}$ times smaller than the ℓ_2 counterpart, showcasing the de-localization effect of the errors of the spectral estimates \mathbf{U} and \mathbf{V} .
- *Entrywise matrix estimation bounds.* Furthermore, the entrywise error (4.43b) is about an order of n_1 times smaller than the corresponding Euclidean error predicted in Theorem 3.23. This indicates that no entry in the resulting matrix estimate suffers from an error significantly higher than the average entrywise error.

4.5 Application: Exact community recovery

Another application that benefits remarkably from the fine-grained eigenvector perturbation theory is community recovery. This section reexplores the stochastic block model studied in Section 3.4, and develops significantly enhanced theoretical support for spectral clustering.

4.5.1 Performance guarantees: Exact recovery

The focus of this section is simultaneous recovery of the community memberships of all vertices, which is termed *exact recovery* or *strong consistency* in the community detection literature (Abbe, 2017). This imposes a much stronger requirement than the *weak consistency* studied in Section 3.4.3.

For the sake of conciseness, the theorem below concentrates on the challenging regime where $p, q \asymp \frac{\log n}{n}$, corresponding to the lowest possible edge densities that allow for exact recovery. This is because, if $p < \frac{\log n}{n}$, then with high probability, one can find isolated vertices that are not connected with any edge in the graph (Durrett, 2007); hence, there will be absolutely no means to infer the community membership of these isolated vertices. The theoretical guarantee is as follows.

Theorem 4.6. Fix any constant $\varepsilon > 0$, and consider the setting of Section 3.4.1. Suppose $p = \frac{\alpha \log n}{n}$ and $q = \frac{\beta \log n}{n}$ for some sufficiently large constants $\alpha > \beta > 0$.¹ In addition, assume that

$$(\sqrt{p} - \sqrt{q})^2 \geq 2(1 + \varepsilon) \frac{\log n}{n}. \quad (4.45)$$

With probability $1 - o(1)$, the spectral method in Section 3.4.2 yields

$$x_i = x_i^* \text{ for all } 1 \leq i \leq n, \quad \text{or} \quad x_i = -x_i^* \text{ for all } 1 \leq i \leq n.$$

This theorem, which first appeared in Abbe *et al.* (2020b), identifies a sufficient recovery condition in terms of the edge densities. The result substantially strengthens the ℓ_2 -based theory in Section 3.4.3, uncovering the capability of the spectral method in achieving not merely almost exact recovery in the average sense, but more appealingly, exact community recovery that ensures correct labels of all vertices.

A natural question arises as to whether the recovery condition (4.45) is improvable via more sophisticated algorithms. Answering this question requires information-theoretic thinking, that is, how to characterize a fundamental threshold—in terms of the difference of edge densities—below which exact recovery is deemed infeasible. As has been demonstrated in Abbe *et al.* (2016), Mossel *et al.* (2015), and Hajek *et al.* (2016), no algorithm whatsoever is able to achieve exact community recovery if

$$(\sqrt{p} - \sqrt{q})^2 \leq 2(1 - \varepsilon) \frac{\log n}{n} \quad (4.46)$$

for any constant $\varepsilon > 0$. This fundamental lower bound, in conjunction with Theorem 4.6, reveals a sharp phase transition behind the performance of the spectral method. In particular, its optimality is guaranteed

¹In the current proof, the constants α, β might depend on the fixed choice of ε . Encouragingly, this restriction can be lifted; see Abbe *et al.* (2020b) for details.

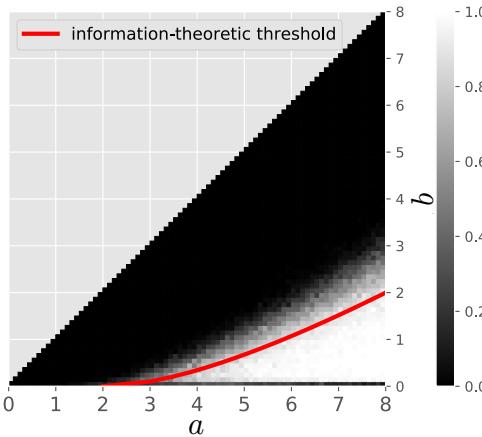


Figure 4.2: Phase transition of spectral methods for exact community recovery. Set $n = 300$, $p = (a \log n)/n$, and $q = (b \log n)/n$. We vary a, b from 0 to 8 with an equal space of 0.1. For each configuration of $a \geq b$, we conduct 100 Monte Carlo trials and report the empirical success rate for recovering the entire community structure correctly. The empirical phase transition occurs near the information-theoretic threshold (see (4.46)).

all the way down to the information-theoretic threshold; see Figure 4.2 for numerical evidence.

Given that the above information-theoretic threshold is specified in terms of $(\sqrt{p} - \sqrt{q})^2$, the reader might naturally wonder what the operational meaning of this quantity is. As it turns out, this metric is a sort of distance measure between the two edge probability distributions under consideration. In truth, in the setting of Theorem 4.6, this metric is intimately related to the squared Hellinger distance between two Bernoulli distributions.

Definition 4.2 (Squared Hellinger distance). Consider two distributions P and Q over a finite alphabet \mathcal{Y} . The squared Hellinger distance $\mathsf{H}^2(P \parallel Q)$ between P and Q is defined as follows

$$\mathsf{H}^2(P \parallel Q) := \frac{1}{2} \sum_{y \in \mathcal{Y}} \left(\sqrt{P(y)} - \sqrt{Q(y)} \right)^2. \quad (4.47)$$

In particular, consider the squared Hellinger distance between two Bernoulli distributions of interest $\text{Bern}(p)$ and $\text{Bern}(q)$, where we denote

by $\text{Bern}(p)$ the Bernoulli distribution with mean p . It is seen that (Chen *et al.*, 2016b)

$$\begin{aligned}\mathsf{H}^2(\text{Bern}(p), \text{Bern}(q)) &:= \frac{1}{2}(\sqrt{p} - \sqrt{q})^2 + \frac{1}{2}(\sqrt{1-p} - \sqrt{1-q})^2 \\ &= (1 + o(1))\frac{1}{2}(\sqrt{p} - \sqrt{q})^2,\end{aligned}$$

when $p = o(1)$ and $q = o(1)$.² The phase transition phenomenon identified in (4.45) and (4.46) can then be alternatively described as

$$\begin{aligned}\text{spectral method works} &\quad \text{if } \mathsf{H}^2(\text{Bern}(p), \text{Bern}(q)) \geq (1 + \varepsilon)\frac{\log n}{n} \\ \text{no algorithm works} &\quad \text{if } \mathsf{H}^2(\text{Bern}(p), \text{Bern}(q)) \leq (1 - \varepsilon)\frac{\log n}{n}\end{aligned}$$

for an arbitrary small constant $\varepsilon > 0$.

4.5.2 Proof of Theorem 4.6

We now turn to the proof of Theorem 4.6. Without loss of generality, suppose that $x_i^* = 1$ for all $1 \leq i \leq n/2$ and $x_i^* = -1$ for all $i > n/2$, so that $\mathbf{u}^* = \frac{1}{\sqrt{n}} \begin{bmatrix} \mathbf{1}_{n/2} \\ -\mathbf{1}_{n/2} \end{bmatrix}$.

Recalling the matrix \mathbf{M} given in (3.30) and its mean \mathbf{M}^* in (3.31), one can immediately see that $\kappa = \mu = r = 1$ for \mathbf{M}^* in this application. Theorem 4.2 (cf. (4.31b)) readily implies the existence of some $z \in \{1, -1\}$ such that

$$\left\| z\mathbf{u} - \frac{1}{\lambda^*} \mathbf{M}\mathbf{u}^* \right\|_\infty \lesssim \frac{\sigma}{|\lambda^*|} + \frac{\sigma^2 \sqrt{n \log n} + \sigma B \log^{3/2} n}{(\lambda^*)^2} \quad (4.48)$$

with probability at least $1 - O(n^{-5})$. Additionally, it has already been explained in Section 3.4 that

$$B = 1, \quad \sigma^2 \leq \max\{p, q\} = p, \quad \text{and} \quad \lambda^* = n(p - q)/2.$$

²To justify this approximation, the following calculation suffices:

$\sqrt{1-q} - \sqrt{1-p} = \frac{p-q}{\sqrt{1-p} + \sqrt{1-q}} = (1 + o(1))(\sqrt{p} - \sqrt{q})(\sqrt{p} + \sqrt{q}) = o(\sqrt{p} - \sqrt{q}).$

Substitution into (4.48) reveals that

$$\begin{aligned}\|z\lambda^* \mathbf{u} - \mathbf{M}\mathbf{u}^*\|_\infty &\lesssim \sigma + \frac{\sigma^2 \sqrt{n \log n}}{\lambda^*} + \frac{\sigma B \log^{3/2} n}{\lambda^*} \\ &\leq C \left(\sqrt{p} + \frac{p \sqrt{\log n}}{\sqrt{n}(p-q)} + \frac{\sqrt{p} \log^{3/2} n}{n(p-q)} \right)\end{aligned}\quad (4.49)$$

holds for some universal constant $C > 0$. As a result, a crucial step boils down to controlling $\mathbf{M}\mathbf{u}^*$ in an entrywise manner: each element is a difference between two independent random binomial random variables and is accomplished through the following lemma.

Lemma 4.7. Suppose that

$$H_{p,q}^2 := (\sqrt{p} - \sqrt{q})^2 \geq (1 + \varepsilon) \frac{2 \log n}{n} \quad (4.50)$$

for some quantity $\varepsilon > 0$. Let $\varepsilon_0 := \frac{\varepsilon \log n}{\sqrt{n} \log \frac{p(1-q)}{q(1-p)}} - \frac{1}{\sqrt{n}}$. Then with probability exceeding $1 - n^{-\varepsilon/2}$, one has

$$\mathbf{M}_{l,\cdot} \mathbf{u}^* \geq \varepsilon_0 \text{ for all } l \leq \frac{n}{2}, \quad \text{and} \quad \mathbf{M}_{l,\cdot} \mathbf{u}^* \leq -\varepsilon_0 \text{ for all } l > \frac{n}{2}.$$

We now return to analyze the entrywise behavior of \mathbf{u} . Note that

$$\begin{aligned}(z\lambda^*)u_l &\geq \mathbf{M}_{l,\cdot} \mathbf{u}^* - |z\lambda^* u_l - \mathbf{M}_{l,\cdot} \mathbf{u}^*| \quad \text{for all } l \leq n/2; \\ (z\lambda^*)u_l &\leq \mathbf{M}_{l,\cdot} \mathbf{u}^* + |z\lambda^* u_l - \mathbf{M}_{l,\cdot} \mathbf{u}^*| \quad \text{for all } l > n/2.\end{aligned}$$

This together with (4.49), Lemma 4.7 and $\lambda^* > 0$ yields that if

$$\frac{\varepsilon \log n}{\sqrt{n} \log \frac{p(1-q)}{q(1-p)}} > \frac{1}{\sqrt{n}} + C \left(\sqrt{p} + \frac{p \sqrt{\log n}}{\sqrt{n}(p-q)} + \frac{\sqrt{p} \log^{3/2} n}{n(p-q)} \right), \quad (4.51)$$

then it follows that

$$zu_l > 0 \quad \text{for all } 1 \leq l \leq \frac{n}{2}, \quad \text{and} \quad zu_l < 0 \quad \text{for all } l > \frac{n}{2},$$

thus guaranteeing exact community recovery once the rounding procedure (based on the sign) is applied.

To finish up, it remains to validate Condition (4.51). Fixing $\varepsilon > 0$ to be a constant, we make the following observations.

- From the assumptions $\varepsilon \asymp 1$ and $p, q \asymp \frac{\log n}{n}$ (or $\alpha, \beta \asymp 1$), one has

$$\frac{\varepsilon \log n}{\sqrt{n} \log \frac{p(1-q)}{q(1-p)}} = \frac{\varepsilon \log n}{\sqrt{n} \log \frac{(1+o(1))\alpha}{\beta}} \asymp \frac{\log n}{\sqrt{n}} \gg \sqrt{p} + \frac{1}{\sqrt{n}}.$$

- Turning to the term $\frac{p\sqrt{\log n}}{\sqrt{n}(p-q)}$, we observe that

$$\begin{aligned} \log \frac{p(1-q)}{q(1-p)} &= \log \left(1 + \frac{p-q}{q(1-p)} \right) \leq \frac{p-q}{q(1-p)} \leq \frac{2(\alpha-\beta)}{\beta}, \\ \implies \frac{\varepsilon \log n}{\sqrt{n} \log \frac{p(1-q)}{q(1-p)}} &\geq \frac{\varepsilon \log n}{2\sqrt{n}} \frac{\beta}{\alpha-\beta}, \end{aligned} \quad (4.52)$$

where in the last inequality of the first line we have used the assumption $p = o(1)$ and hence $1/(1-p) \leq 2$. Given that $\varepsilon, \alpha, \beta \asymp 1$, it is guaranteed that

$$(4.52) \gg \frac{\alpha \sqrt{\log n}}{\sqrt{n}(\alpha-\beta)} = \frac{p \sqrt{\log n}}{\sqrt{n}(p-q)}.$$

- We then move on to the term $\frac{\sqrt{p} \log^{3/2} n}{n(p-q)}$. If $\alpha/\beta \leq 2$ and $\beta \geq \frac{200C^2}{\varepsilon^2} \geq \frac{100C^2\alpha/\beta}{\varepsilon^2}$, then one has $\beta \geq \frac{10C\sqrt{\alpha}}{\varepsilon}$ and hence

$$(4.52) \geq \frac{5C\sqrt{\alpha} \log n}{\sqrt{n}(\alpha-\beta)} = \frac{5C\sqrt{p} \log^{3/2} n}{n(p-q)}.$$

In addition, if $\alpha/\beta > 2$, then it follows that

$$\frac{\sqrt{p} \log^{3/2} n}{n(p-q)} = \frac{\sqrt{\alpha} \log n}{\sqrt{n}(\alpha-\beta)} \leq \frac{2\sqrt{\alpha} \log n}{\sqrt{n}\alpha} = \frac{2 \log n}{\sqrt{n}\alpha}, \quad (4.53)$$

where the inequality holds true since $\alpha-\beta > \alpha-\alpha/2 = \alpha/2$. Using the basic inequality $\log x \leq \sqrt{x}$ further leads to

$$\frac{\varepsilon \log n}{\sqrt{n} \log \frac{p(1-q)}{q(1-p)}} \geq \frac{\varepsilon \log n}{\sqrt{n} \log \frac{2\alpha}{\beta}} \geq \frac{\varepsilon \sqrt{\beta} \log n}{\sqrt{n} \sqrt{2\alpha}} \geq \frac{5C\sqrt{p} \log^{3/2} n}{n(p-q)}.$$

Here, the first inequality holds since $p, q = o(1)$ and hence $\frac{1-q}{1-p} \leq 2$, whereas the last relation relies on (4.53) and holds with the proviso that $\beta \geq 200C^2/\varepsilon^2$.

The above calculations taken collectively establish Condition (4.51) under the assumptions of Theorem 4.6, thus concluding the proof.

Remark 4.2. It is worth pointing out that the bound (4.31a) in Theorem 4.2 is not sufficiently tight when establishing this result. Instead, one needs to resort to the more refined bound (4.31b) in Theorem 4.2, which allows us to sharpen the error bound by explicitly accounting for the first-order error term $(M - M^*)u^*$.

4.5.3 Proof of auxiliary lemmas

Before embarking on the proof of Lemma 4.7, we first record non-asymptotic tail bounds concerning log-likelihood ratios and a sum of Bernoulli random variables, which make apparent the role of the squared Hellinger distance (Tsybakov, 2009).

Lemma 4.8. Consider two distributions P and Q over a finite alphabet \mathcal{Y} , and suppose that $P(y) \neq 0$ for all $y \in \mathcal{Y}$. Generate an independent sequence $\{y_i\}_{1 \leq i \leq n}$ obeying $y_i \sim P$. Then for any $\zeta \in \mathbb{R}$ one has

$$\mathbb{P} \left\{ \sum_{i=1}^n \log \frac{Q(y_i)}{P(y_i)} \geq -n\zeta \right\} \leq \exp \left(-n \left[\mathsf{H}^2(P \parallel Q) - \frac{\zeta}{2} \right] \right), \quad (4.54)$$

where $\mathsf{H}^2(P \parallel Q)$ is the squared Hellinger distance between P and Q defined in (4.47).

Lemma 4.9. Consider two sequences of independent random variables

$$z_i \sim \text{Bern}(p), \quad w_i \sim \text{Bern}(q), \quad 1 \leq i \leq n,$$

and suppose that $p > q$. For any $\xi \in \mathbb{R}$, it follows that

$$\mathbb{P} \left\{ \sum_{i=1}^n (z_i - w_i) \leq n\xi \right\} \leq \exp \left(-n \left[\mathsf{H}_{p,q}^2 - \frac{\xi}{2} \log \frac{p(1-q)}{q(1-p)} \right] \right),$$

where $\mathsf{H}_{p,q}^2 := (\sqrt{p} - \sqrt{q})^2$.

In what follows, we first establish Lemmas 4.8 and 4.9, and then return to prove Lemma 4.7.

Proof of Lemma 4.8. Apply the Chernoff bound to yield

$$\begin{aligned} \mathbb{P} \left\{ \sum_{i=1}^n \log \frac{Q(y_i)}{P(y_i)} \geq -n\zeta \right\} &\leq \frac{\mathbb{E}_{y_i \sim P} \left[\exp \left(\frac{1}{2} \sum_{i=1}^n \log \frac{Q(y_i)}{P(y_i)} \right) \right]}{\exp(-n\zeta/2)} \\ &\stackrel{(i)}{=} \frac{\left(\mathbb{E}_{y \sim P} \left[\left(\frac{Q(y)}{P(y)} \right)^{1/2} \right] \right)^n}{\exp(-n\zeta/2)}, \end{aligned} \quad (4.55)$$

where (i) holds due to the i.i.d. assumption of the y_i 's. In addition,

$$\begin{aligned} \mathbb{E}_{y \sim P} \left[\left(\frac{Q(y)}{P(y)} \right)^{1/2} \right] &= \sum_y P(y) \left(\frac{Q(y)}{P(y)} \right)^{1/2} = \sum_y \sqrt{P(y)Q(y)} \\ &= 1 - \frac{1}{2} \sum_y \left(P(y) + Q(y) - 2\sqrt{P(y)Q(y)} \right) \\ &= 1 - \frac{1}{2} \sum_y \left(\sqrt{P(y)} - \sqrt{Q(y)} \right)^2 \\ &= 1 - \mathsf{H}^2(P \parallel Q) \leq \exp(-\mathsf{H}^2(P \parallel Q)), \end{aligned} \quad (4.56)$$

where the second line follows since $\sum_y P(y) = \sum_y Q(y) = 1$, and the last line uses the definition (4.47) and the elementary inequality $1 - x \leq \exp(-x)$. Substituting (4.56) into (4.55) concludes the proof.

Proof of Lemma 4.9. Set $y_i := z_i - w_i$. The proof is built upon a mapping between $\sum_{i=1}^n y_i$ and a certain log-likelihood ratio. Specifically, let us introduce two distributions P and Q supported on $\{1, 0, -1\}$:

$$\begin{aligned} P(x) &= \begin{cases} p(1-q), & \text{if } x = 1, \\ q(1-p), & \text{if } x = -1, \\ pq + (1-p)(1-q), & \text{if } x = 0, \end{cases} \\ Q(x) &= \begin{cases} q(1-p), & \text{if } x = 1, \\ p(1-q), & \text{if } x = -1, \\ pq + (1-p)(1-q), & \text{if } x = 0. \end{cases} \end{aligned}$$

Apparently, P (resp. Q) corresponds to the distribution of y_i (resp. $-y_i$). A key observation is that

$$\sum_{i=1}^n \log \frac{Q(y_i)}{P(y_i)} = \sum_{i=1}^n \left\{ \mathbb{1}\{y_i = 1\} \log \frac{q(1-p)}{p(1-q)} + \mathbb{1}\{y_i = -1\} \log \frac{p(1-q)}{q(1-p)} \right\}$$

$$= \sum_{i=1}^n y_i \log \frac{q(1-p)}{p(1-q)},$$

which relies on the fact that y_i is supported on $\{1, 0, -1\}$. Recognizing that $\log \frac{p(1-q)}{q(1-p)} > 0$ holds as long as $p > q$ (since $q(1-p) < p(1-q)$), we can further derive

$$\begin{aligned} \mathbb{P} \left\{ \sum_{i=1}^n y_i \leq n\xi \right\} &= \mathbb{P} \left\{ \frac{1}{\log \frac{p(1-q)}{q(1-p)}} \sum_{i=1}^n \log \frac{Q(y_i)}{P(y_i)} \geq -n\xi \right\} \\ &\leq \exp \left(-n \left[\mathbb{H}^2(P \parallel Q) - \frac{\xi}{2} \log \frac{p(1-q)}{q(1-p)} \right] \right), \end{aligned}$$

where the last inequality comes from Lemma 4.8. From the constructions of P and Q and the definition (4.47) of $\mathbb{H}^2(P \parallel Q)$, it is easily seen that

$$\begin{aligned} \mathbb{H}^2(P \parallel Q) &= \left(\sqrt{p(1-q)} - \sqrt{q(1-p)} \right)^2 = \frac{(p-q)^2}{(\sqrt{p(1-q)} + \sqrt{q(1-p)})^2} \\ &\geq \frac{(p-q)^2}{(\sqrt{p} + \sqrt{q})^2} = (\sqrt{p} - \sqrt{q})^2, \end{aligned}$$

thus concluding the proof.

Proof of Lemma 4.7. Let us start by looking at the first entry of $M\mathbf{u}^*$. It is seen from the construction (3.30) that

$$M_{1,\cdot} \mathbf{u}^* = \mathbf{A}_{1,\cdot} \mathbf{u}^* - \frac{p+q}{2} (\mathbf{1}^\top \mathbf{u}^*) \mathbf{1} + pu_1^* \geq \mathbf{A}_{1,\cdot} \mathbf{u}^*, \quad (4.57)$$

where we have used the fact that $\mathbf{1}^\top \mathbf{u}^* = 0$ and $u_1^* > 0$. The expression $\mathbf{u}^* = \frac{1}{\sqrt{n}} \begin{bmatrix} \mathbf{1}_{n/2} \\ -\mathbf{1}_{n/2} \end{bmatrix}$ admits the following decomposition

$$\mathbf{A}_{1,\cdot} \mathbf{u}^* = \frac{1}{\sqrt{n}} \sum_{i=1}^{n/2} (A_{1,i} - A_{1,i+n/2}), \quad (4.58)$$

which can be controlled via Lemma 4.9.

Observe that $A_{1,i} \sim \text{Bern}(p)$ for all $1 < i \leq n/2$ and $A_{1,i} \sim \text{Bern}(q)$ otherwise. Using the definitions of z_i and w_i in Lemma 4.9, we obtain

$$\mathbb{P} \left\{ \sum_{i=1}^{n/2} (A_{1,i} - A_{1,i+n/2}) \leq \frac{n\zeta}{2} - 1 \right\} \leq \mathbb{P} \left\{ \sum_{i=1}^{n/2} (z_i - w_i) \leq \frac{n\zeta}{2} \right\}$$

$$\leq \exp\left(-\frac{n}{2}\left[\mathsf{H}_{p,q}^2 - \frac{\zeta}{2}\log\frac{p(1-q)}{q(1-p)}\right]\right) \leq \frac{1}{n^{1+\delta}} \quad (4.59)$$

for some $\delta > 0$, where the first inequality follows since $A_{1,1} = 0 \leq z_1 + 1$ (so that $A_{1,1} - A_{1,n/2+1}$ is stochastically dominated by $z_1 - w_1$), and the last inequality holds as long as

$$\mathsf{H}_{p,q}^2 - \frac{\zeta}{2}\log\frac{p(1-q)}{q(1-p)} \geq \frac{2(1+\delta)\log n}{n}, \quad (4.60)$$

which we shall ensure at the end of the proof. Substituting (4.59) into (4.58) and (4.57) yields

$$\mathbb{P}\left\{\mathbf{M}_{1,\cdot}\mathbf{u}^* \leq \frac{n\zeta - 2}{2\sqrt{n}}\right\} \leq \mathbb{P}\left\{\mathbf{A}_{1,\cdot}\mathbf{u}^* \leq \frac{n\zeta - 2}{2\sqrt{n}}\right\} \leq \frac{1}{n^{1+\delta}}.$$

Repeating the preceding analysis for $\mathbf{M}_{l,\cdot}\mathbf{u}^*$ with other l 's and taking the union bound, we see that with probability at least $1 - n^{-\delta}$,

$$\mathbf{M}_{l,\cdot}\mathbf{u}^* \geq \frac{n\zeta - 2}{2\sqrt{n}} \quad \text{if } l \leq n/2 \quad (4.61a)$$

$$\mathbf{M}_{l,\cdot}\mathbf{u}^* \leq -\frac{n\zeta - 2}{2\sqrt{n}} \quad \text{if } l > n/2 \quad (4.61b)$$

hold simultaneously for all $1 \leq l \leq n$.

Finally, it remains to ensure satisfaction of (4.60). As it turns out, if the condition (4.50) holds, then it suffices to take $\delta \leq \varepsilon/2$ and $\zeta = \frac{2\varepsilon\log n}{n\log\frac{p(1-q)}{q(1-p)}}$. This completes the proof.

4.6 Distributional theory and uncertainty quantification

Thus far, we have demonstrated intriguing statistical performance of estimators developed based on spectral methods. As one can anticipate, the quality of a spectral estimator is largely affected by the imperfection of data generating mechanisms (e.g., noise corruption, missing data). The uncertainty of the estimator due to these factors would inevitably influence any subsequent decision making based on it. Viewed in this light, it is recommended to accompany the estimator in hand with valid measures of uncertainty (or “confidence”), in order to better inform decision makers.

Take the low-rank matrix estimation problem in Section 4.2.1 for instance: an important uncertainty quantification task can be posed as the construction of a valid confidence interval—based on the spectral estimator—that is likely to cover an unseen entry of the matrix of interest \mathbf{M}^* . More precisely, for any location (i, j) and any target coverage level $1 - \alpha \in (0, 1)$ (e.g., 95%), we aim to identify a short interval—denoted by $\text{Cl}_{i,j}^{1-\alpha}$ —based on the spectral estimator such that

$$\mathbb{P}(M_{i,j}^* \in \text{Cl}_{i,j}^{1-\alpha}) \approx 1 - \alpha, \quad (4.62)$$

which essentially augments a point estimate into an interval that is guaranteed to cover the unknown with the pre-specified target probability. Note that the problem of constructing a valid confidence interval falls within the realm of *statistical inference* in the statistics literature, which constitutes an important step beyond statistical estimation. Accomplishing this task in high dimension often calls for a refined statistical reasoning toolbox that offers quantitative distributional characterizations of the estimator.

4.6.1 Entrywise distributional guarantees

As a natural starting point to build confidence intervals, we seek to develop comprehensive understanding about the distribution of the spectral estimator. In general, obtaining a non-asymptotic yet tractable distributional characterization of a nonconvex estimator like the spectral method could be remarkably challenging. Fortunately, the ℓ_∞ and $\ell_{2,\infty}$ perturbation theory introduced previously (e.g., Theorem 4.2) allows one to make progress for some important scenarios.

Let us revisit the setting in Section 4.2.1, and consider the following estimator of the unknown low-rank matrix \mathbf{M}^* :

$$\widehat{\mathbf{M}} = [\widehat{M}_{i,j}]_{1 \leq i,j \leq n} = \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^\top, \quad (4.63)$$

obtained via the spectral method. The aim is to develop tractable distributional guarantees for each entry of $\widehat{\mathbf{M}} - \mathbf{M}^*$.

Towards this end, we first examine whether our previous results shed light on certain distributional properties of $\widehat{\mathbf{M}} - \mathbf{M}^*$. Informally, Theorem 4.2 (in particular, (4.31b)) reveals that

$$\mathbf{U} \text{sgn}(\mathbf{H}) \approx \mathbf{M} \mathbf{U}^* (\boldsymbol{\Lambda}^*)^{-1} = \mathbf{U}^* + \mathbf{E} \mathbf{U}^* (\boldsymbol{\Lambda}^*)^{-1}. \quad (4.64)$$

Assuming tightness of this first-order approximation, one further derives

$$\begin{aligned}
\mathbf{U}\Lambda\mathbf{U}^\top - \mathbf{M}^* &\stackrel{(i)}{\approx} \mathbf{U}\text{sgn}(\mathbf{H})\Lambda^*(\mathbf{U}\text{sgn}(\mathbf{H}))^\top - \mathbf{U}^*\Lambda^*\mathbf{U}^{*\top} \\
&\stackrel{(ii)}{\approx} (\mathbf{U}\text{sgn}(\mathbf{H}) - \mathbf{U}^*)\Lambda^*\mathbf{U}^{*\top} + \mathbf{U}^*\Lambda^*(\mathbf{U}\text{sgn}(\mathbf{H}) - \mathbf{U}^*)^\top \\
&\stackrel{(iii)}{\approx} \mathbf{E}\mathbf{U}^*(\Lambda^*)^{-1}\Lambda^*\mathbf{U}^{*\top} + \mathbf{U}^*\Lambda^*(\mathbf{E}\mathbf{U}^*(\Lambda^*)^{-1})^\top \\
&= \mathbf{E}\mathbf{U}^*\mathbf{U}^{*\top} + \mathbf{U}^*\mathbf{U}^{*\top}\mathbf{E},
\end{aligned} \tag{4.65}$$

where (i) holds as long as $\text{sgn}(\mathbf{H})\Lambda^*\text{sgn}(\mathbf{H})^\top \approx \Lambda$ (which has already been illuminated in the analysis of Corollary 4.3 and will be solidified momentarily), (ii) is obtained by dropping the higher-order term $(\mathbf{U}\text{sgn}(\mathbf{H}) - \mathbf{U}^*)\Lambda^*(\mathbf{U}\text{sgn}(\mathbf{H}) - \mathbf{U}^*)^\top$, and (iii) relies upon the approximation (4.64).

Given that (4.65) is a linear map of the noise matrix \mathbf{E} , this essentially forms a first-order approximation of $\widehat{\mathbf{M}}$, which in turn enables a tractable distributional theory for $\widehat{\mathbf{M}}$. Observe that each entry of the matrix in (4.65) is a weighted superposition of the independent zero-mean entries of \mathbf{E} . Equipped with this observation, some variant of the central limit theorem suggests that each entry of $\widehat{\mathbf{M}} - \mathbf{M}^*$ is approximately zero-mean Gaussian, as formalized by the theorem below. For notational convenience, we shall define a projection matrix

$$\mathbf{P}^* = [P_{i,j}^*]_{1 \leq i,j \leq n} := \mathbf{U}^*\mathbf{U}^{*\top}, \tag{4.66}$$

and impose a lower bound requirement on the noise variance:

$$\sigma_{\min}^2 \leq \sigma_{i,j}^2 \leq \sigma^2, \quad 1 \leq i, j \leq n. \tag{4.67}$$

Theorem 4.10. Suppose that the assumptions of Theorem 4.2 hold. For any $1 \leq i, j \leq n$, set

$$v_{i,j}^* = \begin{cases} \sum_{l=1}^n \sigma_{i,l}^2 P_{l,j}^{*2} + \sum_{l=1}^n P_{i,l}^{*2} \sigma_{l,j}^2 + 2\sigma_{i,j}^2 P_{i,i}^* P_{j,j}^*, & \text{if } i \neq j, \\ 4 \sum_{l=1}^n \sigma_{i,l}^2 P_{l,i}^{*2}, & \text{if } i = j. \end{cases} \tag{4.68}$$

Assume that $\sigma/\sigma_{\min} = O(1)$, and that

$$\frac{\|\mathbf{U}_{j,\cdot}^*\|_2^2 + \|\mathbf{U}_{i,\cdot}^*\|_2^2}{\|\mathbf{U}^*\|_{\text{F}}^2} \gtrsim \frac{B^2 \kappa^2 \mu^2 r^2 \log^2 n}{\sigma^2 n^2} + \frac{\sigma^2 \mu^2 r \kappa^4 \log^3 n}{(\lambda_r^*)^2}. \tag{4.69}$$

Then the estimator (4.63) obeys

$$\sup_{z \in \mathbb{R}} \left| \mathbb{P} \left(\widehat{M}_{i,j} - M_{i,j}^* \leq z \sqrt{v_{i,j}^*} \right) - \Phi(z) \right| = o(1),$$

where $\Phi(\cdot)$ represents the cumulative density function (CDF) of the standard Gaussian distribution.

The proof of this theorem is postponed to Section 4.11. In a nutshell, Theorem 4.10 tells us that $\widehat{\mathbf{M}}$ is a nearly unbiased estimator of the truth \mathbf{M}^* , as long as the signal strength—as captured by $\|\mathbf{U}_{i,\cdot}^*\|_2$ and $\|\mathbf{U}_{j,\cdot}^*\|_2$ when estimating the (i, j) -th entry—is sufficiently large (cf. (4.69)). The resulting estimation error in each entry is well approximated by a zero-mean Gaussian random variable, whose variance can be determined in a tractable fashion. As can be easily verified, the variance $v_{i,j}^*$ is precisely the variance of the (i, j) -th entry of $\mathbf{E}\mathbf{U}^*\mathbf{U}^{*\top} + \mathbf{U}^*\mathbf{U}^{*\top}\mathbf{E}$ (as singled out in (4.65)). The above distributional theory is non-asymptotic, which lends itself well to high-dimensional applications.

4.6.2 Inference and uncertainty quantification

The Gaussian approximation unveiled in Theorem 4.10, which is dictated by a single parameter $v_{i,j}^*$, paves the way for statistical inference and uncertainty quantification tailored to this model. In order to construct a valid confidence interval for each entry of \mathbf{M}^* , everything boils down to identifying an estimator that approximates the variance parameter $v_{i,j}^*$, ideally in a data-driven yet faithful manner.

In view of the variance characterization (4.68), computing $v_{i,j}^*$ requires information about both the noise variances $\{\sigma_{i,j}^2\}_{1 \leq i,j \leq n}$ and the projection matrix \mathbf{P}^* (cf. (4.66)). However, estimating the noise variances is in general statistically infeasible, given that we only have access to a single observation (i.e., $M_{i,j}$) related to each individual variance $\sigma_{i,j}^2$. Fortunately, the variance $v_{i,j}^*$ involves only the summation or equivalently the average of these individual variances, whose stochastic errors will be averaged out. This leads us to the following surrogate

$$\tilde{v}_{i,j} = \begin{cases} \sum_{l=1}^n E_{i,l}^2 P_{l,j}^{*2} + \sum_{l=1}^n P_{i,l}^{*2} E_{l,j}^2 + 2E_{i,j}^2 P_{i,i}^* P_{j,j}^*, & \text{if } i \neq j, \\ 4 \sum_{l=1}^n E_{i,l}^2 P_{l,i}^{*2}, & \text{if } i = j, \end{cases} \quad (4.70)$$

which is clearly an unbiased estimator of $v_{i,j}^*$. Given the statistical independence of $\{E_{i,j}\}_{i,j}$, we can expect to have $\tilde{v}_{i,j} \approx v_{i,j}^*$, owing to the concentration of measure.

However, the above surrogate $\tilde{v}_{i,j}$ remains practically incomputable, due to the absence of knowledge about both \mathbf{E} and \mathbf{P}^* . To address this issue, we propose the following plug-in estimator:

$$\hat{v}_{i,j} = \begin{cases} \sum_{l=1}^n \hat{E}_{i,l}^2 \hat{P}_{l,j}^2 + \sum_{l=1}^n \hat{P}_{i,l}^2 \hat{E}_{l,j}^2 + 2\hat{E}_{i,j}^2 \hat{P}_{i,i} \hat{P}_{j,j}, & \text{if } i \neq j, \\ 4 \sum_{l=1}^n \hat{E}_{i,l}^2 \hat{P}_{l,i}^2, & \text{if } i = j, \end{cases} \quad (4.71)$$

where $\hat{\mathbf{E}} = [\hat{E}_{i,j}]_{1 \leq i,j \leq n}$ and $\hat{\mathbf{P}} = [\hat{P}_{i,j}]_{1 \leq i,j \leq n}$ stand for estimators of \mathbf{E} and \mathbf{P}^* , respectively. In particular, we employ the following specific estimators of \mathbf{E} and \mathbf{P}^* , again adopting the plug-in strategy:

$$\hat{\mathbf{E}} := \mathbf{M} - \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^\top, \quad (4.72a)$$

$$\hat{\mathbf{P}} := \mathbf{U} \mathbf{U}^\top, \quad (4.72b)$$

where \mathbf{U} and $\boldsymbol{\Lambda}$ are, as usual, computed via eigendecomposition of \mathbf{M} . For a prescribed coverage level $1 - \alpha$ (with $0 < \alpha < 1$), we construct the following confidence interval for the (i, j) -th entry of \mathbf{M}^* , motivated by the Gaussian approximation in Theorem 4.10:

$$\text{CI}_{i,j}^{1-\alpha} := \left[\hat{M}_{i,j} \pm \Phi^{-1}(1 - \alpha/2) \sqrt{\hat{v}_{i,j}} \right]. \quad (4.73)$$

Here and throughout, for any $b > 0$, we let $[a \pm b]$ abbreviate the interval $[a - b, a + b]$, and we use $\Phi^{-1}(\cdot)$ to represent the inverse CDF of the standard Gaussian distribution.

As encouraging news, the above construction of entrywise confidence intervals is provably valid with high probability, as revealed by the following theorem. The proof is postponed to Section 4.12.

Theorem 4.11. Consider the settings and assumptions in Section 4.2.1, and suppose that $\sigma/\sigma_{\min} = O(1)$, $\kappa^4 \mu^2 r^2 \log n \leq n$ and $\sigma \kappa \sqrt{n \log n} \lesssim |\lambda_r^*|$. Consider any $1 \leq i, j \leq n$, and assume that

$$\frac{\|\mathbf{U}_{j,\cdot}^*\|_2^2 + \|\mathbf{U}_{i,\cdot}^*\|_2^2}{\|\mathbf{U}^*\|_{\text{F}}^2} \gtrsim \frac{B \kappa^2 \mu^2 r^2 \log^2 n}{\sigma n^{3/2}} + \frac{\sigma \mu^2 r \kappa^3 \log^3 n}{|\lambda_r^*| \sqrt{n}}. \quad (4.74)$$

For any fixed coverage level $1 - \alpha \in (0, 1)$, the confidence interval $\text{CI}_{i,j}^{1-\alpha}$ constructed in (4.73) obeys

$$\mathbb{P}\left(M_{i,j}^* \in \text{CI}_{i,j}^{1-\alpha}\right) = 1 - \alpha + o(1). \quad (4.75)$$

Theorem 4.11 confirms that the confidence interval proposed above meets the prescribed coverage requirement, provided that the associated signal strength is not too low (see (4.74)). In addition to its statistical validity, the proposed procedure enjoys several features that make it practically appealing:

- *Adaptive to unknown noise levels and distributions.* The above inference procedure is fully data-driven, which does not require prior knowledge about the noise levels or noise distributions. As alluded to previously, it is in general impossible to estimate the noise variance in each entry, and hence a data-driven yet valid approach is of critical value.
- *Adaptive to heteroskedastic noise.* Our statistical guarantees hold without relying on homogeneity of noise components. In other words, this inference procedure automatically accommodates heteroskedastic noise, a scenario where the variance of the noise components might vary across different locations.

Careful readers might remark that Theorem 4.11 is concerned with statistical inference for a single entry. Interestingly, the distributional theory presented in Section 4.6.1 (see also Lemma 4.18 in the proof of Theorem 4.10) might also be instrumental in pursuing simultaneous inference, namely, the problem of constructing a confidence region that simultaneously accounts for more than one unknown entries. We omit such an extension for the sake of conciseness.

4.7 Application: Confidence intervals for matrix completion

As an illustration of the applicability of the inference procedure described in Section 4.6.2, we develop concrete consequences of Theorem 4.11 in application to noisy matrix completion—an extension of the formulation in Section 3.8 to noisy settings.

Model: noisy matrix completion. Suppose that we are asked to reconstruct a symmetric rank- r matrix $\mathbf{M}^* = [M_{i,j}^*]_{1 \leq i,j \leq n} \in \mathbb{R}^{n \times n}$ with eigendecomposition $\mathbf{M}^* = \mathbf{U}^* \boldsymbol{\Lambda}^* \mathbf{U}^{*\top}$. We only get to acquire noisy observations of a subset of the entries of \mathbf{M}^* ; more precisely, there exists a sampling set $\Omega \subseteq [n] \times [n]$ such that we observe

$$M_{k,l}^* + \eta_{k,l} \quad \text{if } (k, l) \in \Omega. \quad (4.76)$$

Here, $\{\eta_{k,l} \mid k \geq l\}$ denotes independent Gaussian noise obeying

$$\eta_{k,l} = \eta_{l,k} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_\eta^2), \quad k \geq l. \quad (4.77)$$

As before, we focus on the random sampling model such that each location (k, l) with $k \geq l$ is included in the sampling set Ω independently with probability p . Further, assume that \mathbf{M}^* has eigenvalues obeying (4.22), condition number κ (cf. (4.23)), and incoherence parameter μ (cf. (4.24)). Can we build a confidence interval for each entry $M_{i,j}^*$, on the basis of the output of the spectral method?

Computing entrywise confidence intervals. In order to apply the inference procedure in Section 4.6.2, it suffices to determine the data matrix $\mathbf{M} = [M_{i,j}]_{1 \leq i,j \leq n}$, which can be selected as usual. Specifically, a possible inference procedure proceeds as follows:

- Set \mathbf{M} such that for any $1 \leq i, j \leq n$,

$$M_{i,j} = \begin{cases} \frac{1}{p}(M_{i,j}^* + \eta_{i,j}), & \text{if } (i, j) \in \Omega, \\ 0, & \text{else,} \end{cases} \quad (4.78)$$

which clearly obeys $\mathbb{E}[\mathbf{M}] = \mathbf{M}^*$.

- Compute the estimate $\widehat{\mathbf{M}}$ (cf. (4.63)) via the spectral method.
- For a given coverage level $1 - \alpha$ and a given pair (i, j) , construct the confidence interval $\text{CI}_{i,j}^{1-\alpha}$ according to (4.73), with auxiliary parameters provided in (4.71) and (4.72).

Performance guarantees and implications. When specialized to noisy matrix completion, our inference theory in Theorem 4.11 leads to the following statistical guarantees.

Theorem 4.12. Consider the noisy matrix completion setting in this section. Suppose that $\kappa^4 \mu^2 r^2 \log n \leq n$, $\frac{\max_{k,l} |M_{k,l}^*|}{\min_{k,l} |M_{k,l}^*|} = O(1)$,

$$np \gtrsim \kappa^4 r \log n \quad \text{and} \quad \sigma_\eta \kappa \sqrt{\frac{n \log n}{p}} \lesssim |\lambda_r^*|. \quad (4.79)$$

Consider any $1 \leq i, j \leq n$, and assume that

$$\frac{\|U_{j,\cdot}^*\|_2^2 + \|U_{i,\cdot}^*\|_2^2}{\|U^*\|_F^2} \gtrsim \frac{\mu^2 r^2 \kappa^4 \log^3 n}{n \sqrt{np}} + \frac{\sigma_\eta \mu^2 r \kappa^3 \log^3 n}{|\lambda_r^*| \sqrt{np}}. \quad (4.80)$$

For any fixed coverage level $1 - \alpha \in (0, 1)$, the confidence interval $\text{CI}_{i,j}^{1-\alpha}$ constructed in (4.73) obeys

$$\mathbb{P}\left(M_{i,j}^* \in \text{CI}_{i,j}^{1-\alpha}\right) = 1 - \alpha + o(1). \quad (4.81)$$

In order to help interpret the applicable range of Theorem 4.12, let us focus on the simple scenario with $\kappa, \mu, r \asymp 1$ to simplify discussion.

- First of all, Condition (4.79) can be simplified as

$$np \gtrsim \log n \quad \text{and} \quad \sigma_\eta \sqrt{\frac{n \log n}{p}} \lesssim |\lambda_r^*|.$$

The first condition on the sampling size coincides with the fundamental requirement even if the goal is merely to enable reliable estimation (Candès and Tao, 2010), whereas the second condition on the signal-to-noise ratio is also necessary—up to some log factor—to ensure an estimation quality better than that of a random guess (Cai *et al.*, 2021a, Theorem 3.3).

- Next, we move on to interpret the other condition (4.80) imposed in our theory, which simplifies to

$$\frac{\|U_{j,\cdot}^*\|_2^2 + \|U_{i,\cdot}^*\|_2^2}{\|U^*\|_F^2} \gtrsim \frac{\log^3 n}{n \sqrt{np}} + \frac{\sigma_\eta \log^3 n}{|\lambda_r^*| \sqrt{np}}.$$

Let us consider the most challenging case where $np \gtrsim \text{poly log } n$ and $\sigma_\eta \sqrt{\frac{n \text{poly log } n}{p}} \lesssim |\lambda_r^*|$ (for some sufficiently large poly-log factor). In such a case, the above condition only requires

$$\frac{\|\mathbf{U}_{j,\cdot}^*\|_2^2 + \|\mathbf{U}_{i,\cdot}^*\|_2^2}{\|\mathbf{U}^*\|_{\text{F}}^2} \gtrsim \frac{1}{n \text{poly log } n},$$

indicating that the associated signal power $\|\mathbf{U}_{j,\cdot}^*\|_2^2 + \|\mathbf{U}_{i,\cdot}^*\|_2^2$ is allowed to be much smaller than the average signal power across all rows (which can be captured by $\|\mathbf{U}^*\|_{\text{F}}^2/n$).

In a nutshell, the validity of our inference procedure is ensured for broad settings. Additionally, we have conducted a series of numerical experiments to examine the entrywise distributions of \mathbf{M} . As illustrated in Figure 4.3, the normalized estimation error $(\widehat{v}_{i,j})^{-1/2}(\widehat{M}_{i,j} - M_{i,j}^*)$ is close in distribution to a standard Gaussian random variable, which corroborates our theory on the confidence interval construction.

Before concluding, we would like to remark that: while the distributional theory for spectral methods allows for valid construction of confidence intervals for an unseen entry, it is oftentimes not among the most effective statistical inference procedures that one can put forward. There exist other alternatives that are provably more efficient, including but not limited to inference procedures based on convex relaxation and nonconvex optimization (Chen *et al.*, 2019c; Xia and Yuan, 2021), and the ones based on more refined spectral methods (Yan *et al.*, 2021; Chernozhukov *et al.*, 2021).

Proof of Theorem 4.12. Given that $\eta_{i,j}$ is a Gaussian random variable and hence possibly unbounded, we find it convenient to introduce a truncated version as follows

$$\tilde{\eta}_{i,j} = \eta_{i,j} \mathbb{1}\{|\eta_{i,j}| \leq 5\sigma_\eta \sqrt{\log n}\}, \quad 1 \leq i, j \leq n.$$

and

$$\widetilde{M}_{i,j} = \begin{cases} \frac{1}{p}(M_{i,j}^* + \tilde{\eta}_{i,j}), & \text{if } (i, j) \in \Omega, \\ 0, & \text{else.} \end{cases}$$

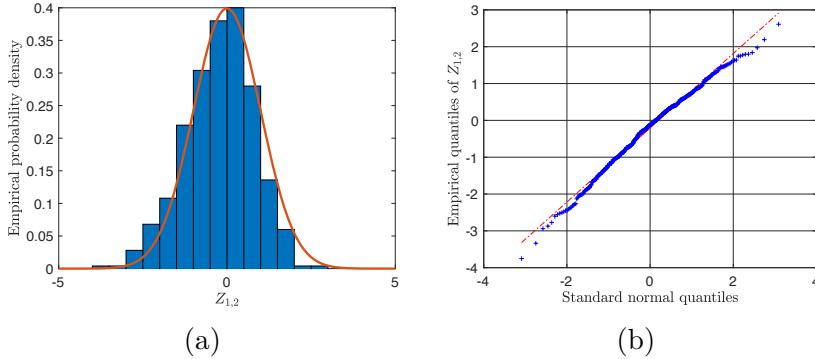


Figure 4.3: Entrywise numerical distribution for noisy matrix completion. We generate $\mathbf{M}^* = \mathbf{U}^* \mathbf{U}^{*\top}$ with \mathbf{U}^* being a random orthonormal matrix, and analyze the behavior of $Z_{1,2} = (\widehat{v}_{1,2})^{-0.5} (\widehat{M}_{1,2} - M_{1,2}^*)$, where \widehat{M} is defined in (4.63) and $\widehat{v}_{i,j}$ is defined in (4.71). The results are reported for 500 Monte Carlo trials when $p = 0.3$, $n = 1000$, $\sigma_\eta = 10^{-4}$, and $r = 3$. (a) Histogram of the empirical distribution of $Z_{1,2}$; (b) Q-Q (quantile-quantile) plot of $Z_{1,2}$ vs. the standard normal distribution.

Repeating the analysis in Section 3.2.3, we can show that

$$\mathbb{P}\{\mathbf{M} = \widetilde{\mathbf{M}}\} = \mathbb{P}\{\eta_{i,j} = \widetilde{\eta}_{i,j}, \forall i, j \in [n]\} \geq 1 - n^{-10},$$

meaning that \mathbf{M} and $\widetilde{\mathbf{M}}$ are equivalent with high probability. As a result, we shall concentrate on validating the confidence interval computed based on $\widetilde{\mathbf{M}}$ in the subsequent analysis. Before proceeding, we record several key properties about $\widetilde{\eta}_{i,j}$ as follows:

$$\mathbb{E}[\widetilde{\eta}_{i,j}] = 0, \quad \mathbb{E}[\widetilde{\eta}_{i,j}^2] = (1 - o(1))\sigma_\eta^2, \quad |\widetilde{\eta}_{i,j}| \leq 5\sigma_\eta\sqrt{\log n}. \quad (4.82)$$

The proof follows by invoking Theorem 4.11, as long as the conditions required therein are satisfied. To begin with, the associated variance parameters are given by

$$\begin{aligned} \sigma_{i,j}^2 &:= \mathbb{E}[(\widetilde{M}_{i,j} - M_{i,j}^*)^2] \\ &= p\mathbb{E}\left[\left(\frac{1-p}{p}M_{i,j}^* + \frac{1}{p}\widetilde{\eta}_{i,j}\right)^2\right] + (1-p)(M_{i,j}^*)^2 \\ &= \frac{(1-p)^2}{p}(M_{i,j}^*)^2 + \frac{1}{p}\mathbb{E}[\widetilde{\eta}_{i,j}^2] + (1-p)(M_{i,j}^*)^2 \end{aligned}$$

$$= \frac{1-p}{p} (M_{i,j}^*)^2 + \frac{1-o(1)}{p} \sigma_\eta^2,$$

thus leading to

$$\begin{aligned} \sigma_{\min}^2 &:= \min_{i,j} \sigma_{i,j}^2 = \frac{1-p}{p} \min_{i,j} (M_{i,j}^*)^2 + \frac{1-o(1)}{p} \sigma_\eta^2, \\ \sigma^2 &:= \max_{i,j} \sigma_{i,j}^2 = \frac{1-p}{p} \|\mathbf{M}^*\|_\infty^2 + \frac{1-o(1)}{p} \sigma_\eta^2. \end{aligned}$$

Apparently, $\sigma^2/\sigma_{\min}^2 = O(1)$ holds true under the assumptions of Theorem 4.12. In addition, the random variables $\{\widetilde{M}_{i,j} - M_{i,j}^*\}$ are all bounded obeying

$$|\widetilde{M}_{i,j} - M_{i,j}^*| \leq \frac{1-p}{p} |M_{i,j}^*| + \frac{1}{p} |\widetilde{\eta}_{i,j}| \leq \frac{(1-p) \|\mathbf{M}^*\|_\infty + \sigma_\eta \sqrt{5 \log n}}{p} =: B.$$

These bounds readily imply that

$$\frac{B}{\sigma} \asymp \frac{\frac{(1-p) \|\mathbf{M}^*\|_\infty + \sigma_\eta \sqrt{5 \log n}}{p}}{\sqrt{\frac{1-p}{p} \|\mathbf{M}^*\|_\infty^2 + \frac{1}{\sqrt{p}} \sigma_\eta^2}} \lesssim \sqrt{\frac{\log n}{p}}. \quad (4.83)$$

Moving to the condition $\sigma \kappa \sqrt{n \log n} \lesssim |\lambda_r^*|$ in Theorem 4.12, it can be guaranteed if

$$\|\mathbf{M}^*\|_\infty \kappa \sqrt{\frac{(1-p)n \log n}{p}} \lesssim |\lambda_r^*| \quad \text{and} \quad \sigma_\eta \kappa \sqrt{\frac{n \log n}{p}} \lesssim |\lambda_r^*|.$$

Given the assumption $\frac{\max_{k,l} |M_{k,l}^*|}{\min_{k,l} |M_{k,l}^*|} = O(1)$, one has

$$\|\mathbf{M}^*\|_\infty \asymp \frac{1}{n} \|\mathbf{M}^*\|_{\text{F}} \leq \frac{\sqrt{r}}{n} \|\mathbf{M}^*\| = \frac{\kappa \sqrt{r}}{n} |\lambda_r^*. \quad (4.84)$$

As a consequence, the condition $\sigma \kappa \sqrt{n \log n} \lesssim |\lambda_r^*|$ can be ensured under Condition (4.79).

It remains to certify Condition (4.74). By virtue of the above calculations of σ and B as well as the property (4.83), it is easily seen that Condition (4.74) is valid as long as the following holds:

$$\frac{\|\mathbf{U}_{j,\cdot}^*\|_2^2 + \|\mathbf{U}_{i,\cdot}^*\|_2^2}{\|\mathbf{U}^*\|_{\text{F}}^2} \gtrsim \frac{\kappa^2 \mu^2 r^2 \log^{5/2} n}{n \sqrt{np}} + \frac{\sqrt{1-p} \|\mathbf{M}^*\|_\infty \mu^2 r \kappa^3 \log^3 n}{|\lambda_r^*| \sqrt{np}}$$

$$+ \frac{\sigma_\eta \mu^2 r \kappa^3 \log^3 n}{|\lambda_r^*| \sqrt{np}}.$$

Taking this together with the relation (4.84), we can demonstrate straightforwardly that Condition (4.74) is guaranteed to hold as long as Condition (4.80) is satisfied. This completes the proof.

4.8 Appendix A: Proof of Theorem 4.2

To simplify notation, we assume throughout the proof that $\lambda_r^* > 0$, namely,

$$|\lambda_1^*| \geq \cdots \geq |\lambda_{r-1}^*| \geq \lambda_r^* > 0. \quad (4.85)$$

The challenge of the proof arises due to the complicated statistical dependency between \mathbf{M} and \mathbf{U} , and the leave-one-out analysis paves a plausible path to decouple the dependency.

4.8.1 Construction of leave-one-out auxiliary estimates

As elucidated in the rank-1 matrix denoising example in Section 4.1, the key to enabling fine-grained analysis is to seek assistance from a collection of leave-one-out estimates. Akin to Section 4.1.3, for each $1 \leq l \leq n$, we construct two auxiliary matrices $\mathbf{M}^{(l)}$ and $\mathbf{E}^{(l)} = [E_{i,j}^{(l)}]_{1 \leq i,j \leq n}$ as follows:

$$\mathbf{M}^{(l)} := \mathbf{M}^* + \mathbf{E}^{(l)}, \quad E_{i,j}^{(l)} := \begin{cases} E_{i,j}, & \text{if } i \neq l \text{ and } j \neq l, \\ 0, & \text{else,} \end{cases} \quad (4.86)$$

which are generated by simply discarding all random noise incurred in the l -th column/row of the data matrix. In addition, let $\lambda_1^{(l)}, \dots, \lambda_n^{(l)}$ be the eigenvalues of $\mathbf{M}^{(l)}$ sorted by

$$|\lambda_1^{(l)}| \geq |\lambda_2^{(l)}| \geq \cdots \geq |\lambda_n^{(l)}|, \quad (4.87)$$

and denote by $\mathbf{u}_i^{(l)}$ the eigenvector of $\mathbf{M}^{(l)}$ associated with $\lambda_i^{(l)}$. The leave-one-out spectral estimates $\mathbf{U}^{(l)}$ and $\mathbf{\Lambda}^{(l)}$ are, therefore, given by

$$\mathbf{U}^{(l)} := [\mathbf{u}_1^{(l)}, \dots, \mathbf{u}_r^{(l)}] \in \mathbb{R}^{n \times r}; \quad \mathbf{\Lambda}^{(l)} := \text{diag}([\lambda_1^{(l)}, \dots, \lambda_r^{(l)}]). \quad (4.88)$$

We emphasize again that the main advantage of introducing the leave-one-out estimate $\mathbf{U}^{(l)}$ stems from its statistical independence from the l -th row of \mathbf{M} , which substantially simplifies the analysis for the l -th row of the estimate. In principle, our analysis employs the leave-one-out estimates to help decouple delicate statistical dependency in a row-by-row fashion. Another crucial aspect of the analysis lies in the exploitation of the proximity of all these auxiliary estimates, a feature that is enabled by the “stability” of the spectral method.

4.8.2 Preliminary facts

Before embarking on the ℓ_∞ and $\ell_{2,\infty}$ analyses, we gather a couple of useful facts, whose proofs are postponed to Section 4.8.4. In what follows, Θ (resp. $\Theta^{(l)}$) denotes a diagonal matrix whose diagonal entries are the principal angles between \mathbf{U} (resp. $\mathbf{U}^{(l)}$) and \mathbf{U}^* . In addition, we find it helpful to introduce the following matrices

$$\mathbf{H} := \mathbf{U}^\top \mathbf{U}^* \quad \text{and} \quad \mathbf{H}^{(l)} := \mathbf{U}^{(l)\top} \mathbf{U}^*, \quad (4.89)$$

which turn out to be close to being orthonormal.

The first set of results follows from the statistical nature of the perturbation matrix \mathbf{E} (cf. Assumption 4.1), which is immediately available from the matrix tail bounds.

Lemma 4.13. Consider the setting in Section 4.2. There is some constant $c_2 > 0$ such that with probability at least $1 - O(n^{-7})$,

$$\max_l \|\mathbf{E}^{(l)}\| \leq \|\mathbf{E}\| \leq c_2 \sigma \sqrt{n}. \quad (4.90)$$

Moreover, for any fixed matrix $\mathbf{A} \in \mathbb{R}^{n \times d}$ with $d \leq n$, one has

$$\|\mathbf{E}\mathbf{A}\|_{2,\infty} \leq 4\sigma \sqrt{\log n} \|\mathbf{A}\|_{\text{F}} + (6B \log n) \|\mathbf{A}\|_{2,\infty}. \quad (4.91)$$

Remark 4.3. In view of (4.91), with probability at least $1 - 2n^{-5}$,

$$\begin{aligned} \|\mathbf{E}\mathbf{U}^*\|_{2,\infty} &\leq 4\sigma \sqrt{\log n} \|\mathbf{U}^*\|_{\text{F}} + (6B \log n) \|\mathbf{U}^*\|_{2,\infty} \\ &= 4\sigma \sqrt{r \log n} + 6B \sqrt{\frac{\mu r \log^2 n}{n}} = (4 + 6c_b)\sigma \sqrt{r \log n}, \end{aligned} \quad (4.92)$$

which relies on the definition (4.24) and the definition of c_b in (4.29). As a result,

$$\begin{aligned}\|\mathbf{M}\mathbf{U}^*\|_{2,\infty} &\leq \|\mathbf{M}^*\mathbf{U}^*\|_{2,\infty} + \|\mathbf{E}\mathbf{U}^*\|_{2,\infty} \\ &\leq \sqrt{\frac{\mu r}{n}}|\lambda_1^*| + (4 + 6c_b)\sigma\sqrt{r\log n},\end{aligned}\quad (4.93)$$

which follows from the fact $\mathbf{M}^*\mathbf{U}^* = \mathbf{U}^*\mathbf{\Lambda}^*$ (so that $\|\mathbf{M}^*\mathbf{U}^*\|_{2,\infty} \leq \|\mathbf{U}^*\|_{2,\infty}\|\mathbf{\Lambda}^*\| = \sqrt{\mu r/n}|\lambda_1^*|$).

With the size of the perturbations (i.e., $\|\mathbf{E}^{(l)}\|$ and $\|\mathbf{E}\|$) under control, the ℓ_2 perturbation theory established in Chapter 2 leads to the following set of conclusions.

Lemma 4.14. Suppose that $c_2\sigma\sqrt{n} \leq (1 - 1/\sqrt{2})\lambda_r^*$, where c_2 is the same constant as in Lemma 4.13. Then with probability at least $1 - O(n^{-7})$, one has

$$\text{dist}(\mathbf{U}, \mathbf{U}^*) \leq \frac{2c_2\sigma\sqrt{n}}{\lambda_r^*}, \quad \text{dist}(\mathbf{U}^{(l)}, \mathbf{U}^*) \leq \frac{2c_2\sigma\sqrt{n}}{\lambda_r^*}, \quad (4.94a)$$

$$\|\sin \Theta\| \leq \frac{c_2\sigma\sqrt{2n}}{\lambda_r^*}, \quad \|\sin \Theta^{(l)}\| \leq \frac{c_2\sigma\sqrt{2n}}{\lambda_r^*}, \quad (4.94b)$$

$$\max_{1 \leq j \leq r} |\lambda_j| \geq \lambda_r^* - c_2\sigma\sqrt{n}, \quad \max_{1 \leq j \leq r} |\lambda_j^{(l)}| \geq \lambda_r^* - c_2\sigma\sqrt{n}, \quad (4.94c)$$

$$\max_{j:j > r} |\lambda_j| \leq c_2\sigma\sqrt{n}, \quad \max_{j:j > r} |\lambda_j^{(l)}| \leq c_2\sigma\sqrt{n} \quad (4.94d)$$

hold simultaneously for all $1 \leq l \leq n$. In addition,

$$\|\mathbf{U}\mathbf{H} - \mathbf{U}^*\|_{\text{F}} \leq \frac{2c_2\sigma\sqrt{rn}}{\lambda_r^*}. \quad (4.94e)$$

with probability exceeding $1 - 2n^{-5}$.

Remark 4.4. Lemmas 4.13 and 4.14 allow us to bound the eigengap and perturbation size as follows

$$|\lambda_r^{(l)}| - |\lambda_{r+1}^{(l)}| \geq \lambda_r^* - c_2\sigma\sqrt{n} - c_2\sigma\sqrt{n} \geq \lambda_r^*/2, \quad (4.95a)$$

$$\begin{aligned}\|\mathbf{M} - \mathbf{M}^{(l)}\| &\leq \|\mathbf{M} - \mathbf{M}^*\| + \|\mathbf{M}^* - \mathbf{M}^{(l)}\| = \|\mathbf{E}\| + \|\mathbf{E}^{(l)}\| \\ &\leq 2c_2\sigma\sqrt{n} \leq (1 - 1/\sqrt{2})(|\lambda_r^{(l)}| - |\lambda_{r+1}^{(l)}|),\end{aligned}\quad (4.95b)$$

which are valid as long as $20c_2\sigma\sqrt{n} \leq \lambda_r^*$. These will prove useful when bounding the approximation error of \mathbf{U} using $\mathbf{U}^{(l)}$.

Another collection of results is concerned with \mathbf{H} and $\mathbf{H}^{(l)}$.

Lemma 4.15. Suppose that the assumptions of Lemma 4.14 hold. With probability at least $1 - O(n^{-7})$,

$$\|\mathbf{H}^{-1}\| \leq 2, \quad \|(\mathbf{H}^{(l)})^{-1}\| \leq 2, \quad (4.96a)$$

$$\|\mathbf{H} - \text{sgn}(\mathbf{H})\| \leq \frac{2c_2^2 n \sigma^2}{(\lambda_r^*)^2}, \quad \|\mathbf{H}^{(l)} - \text{sgn}(\mathbf{H}^{(l)})\| \leq \frac{2c_2^2 n \sigma^2}{(\lambda_r^*)^2} \quad (4.96b)$$

hold simultaneously for all $1 \leq l \leq n$.

Remark 4.5. As a consequence of (4.96) and Proposition 2.1, one has

$$\|\mathbf{A}\| = \|\mathbf{A}\mathbf{H}\mathbf{H}^{-1}\| \leq \|\mathbf{A}\mathbf{H}\| \|\mathbf{H}^{-1}\| \leq 2\|\mathbf{A}\mathbf{H}\|, \quad (4.97a)$$

$$\|\mathbf{A}\| \leq \|\mathbf{A}\mathbf{H}^{(l)}\| \|(\mathbf{H}^{(l)})^{-1}\| \leq 2\|\mathbf{A}\mathbf{H}^{(l)}\| \quad (4.97b)$$

for any matrix \mathbf{A} . Here, $\|\mathbf{A}\|$ could either be the Frobenius norm or the $\ell_{2,\infty}$ norm $\|\cdot\|_{2,\infty}$.

4.8.3 Leave-one-out analysis

Now we move on to the main part of the analysis, which is further decomposed into four steps.

Step 1: decomposing the $\ell_{2,\infty}$ estimation error of \mathbf{U}

By virtue of the proximity of \mathbf{H} and $\text{sgn}(\mathbf{H})$ unveiled in Lemma 4.15, we are allowed to employ $\mathbf{U}\mathbf{H}$ as a surrogate for $\mathbf{U}\text{sgn}(\mathbf{H})$, which is more convenient to work with. As it turns out, the discrepancy between $\mathbf{U}\mathbf{H}$ and the first-order approximation $\mathbf{M}\mathbf{U}^*(\Lambda^*)^{-1}$, and the discrepancy between $\mathbf{U}\mathbf{H}$ and the truth, can be bounded by three important terms, as asserted below. The proof is built upon elementary algebra and basic ℓ_2 perturbation bounds in Section 4.8.2, and is deferred to Section 4.8.4.

Lemma 4.16. Suppose that $2c_2\sigma\sqrt{n} \leq \lambda_r^*$ for some sufficiently large constant $c_2 > 0$. Then with probability at least $1 - O(n^{-7})$, one has

$$\|\mathbf{U}\mathbf{H} - \mathbf{M}\mathbf{U}^*(\Lambda^*)^{-1}\|_{2,\infty} \leq \mathcal{E}_1 + \mathcal{E}_2, \quad (4.98a)$$

$$\|\mathbf{U}\mathbf{H} - \mathbf{U}^*\|_{2,\infty} \leq \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3, \quad (4.98b)$$

where $\mathcal{E}_1 := \frac{2\|\mathbf{M}(\mathbf{U}\mathbf{H} - \mathbf{U}^*)\|_{2,\infty}}{\lambda_r^*}$, $\mathcal{E}_2 := \frac{4\|\mathbf{M}\mathbf{U}^*\|_{2,\infty}\|\mathbf{E}\|}{(\lambda_r^*)^2}$, and $\mathcal{E}_3 := \frac{\|\mathbf{E}\mathbf{U}^*\|_{2,\infty}}{\lambda_r^*}$.

Lemma 4.16 leaves us with three important terms to deal with. The term \mathcal{E}_1 is most complicated as it involves the product of two random matrices, whereas \mathcal{E}_2 and \mathcal{E}_3 can be controlled straightforwardly through our preliminary facts in Section 4.8.2. Specifically, the term \mathcal{E}_2 can be bounded by combining (4.93) with the bound (4.90) on \mathbf{E} to obtain

$$\mathcal{E}_2 \leq \frac{4c_2\kappa\sigma\sqrt{\mu r}}{\lambda_r^*} + \frac{4c_2(4+6c_b)\sigma^2\sqrt{rn\log n}}{(\lambda_r^*)^2}. \quad (4.99)$$

Regarding the term \mathcal{E}_3 , the inequality (4.92) readily gives

$$\mathcal{E}_3 \leq \frac{(4+6c_b)\sigma\sqrt{r\log n}}{\lambda_r^*}. \quad (4.100)$$

Turning to controlling the remaining term \mathcal{E}_1 , a closer inspection, however, reveals substantial challenges, due to the complicated statistical dependency between \mathbf{M} and \mathbf{U} . To further complicate matters, the term \mathcal{E}_1 —as we shall demonstrate momentarily—depend on some intrinsic properties of interest about \mathbf{U} (e.g., $\|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty}$), which might lead to circular reasoning if not handled properly. In order to circumvent this issue, we intend to establish the following relation

$$\mathcal{E}_1 \leq \mathcal{E}_{1,1} + \rho_1 \|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty} \quad (4.101)$$

for some quantity $\mathcal{E}_{1,1} > 0$ that does not involve $\|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty}$ as well as some contraction factor $0 < \rho_1 \leq 1/2$. Assuming the relation (4.101) holds for the moment, we have the following useful claim (the proof is straightforward and again postponed to Section 4.8.4).

Lemma 4.17. If Conditions (4.98) and (4.101) hold with $0 < \rho_1 \leq 1/2$, then we have

$$\|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty} \leq 2(\mathcal{E}_{1,1} + \mathcal{E}_2 + \mathcal{E}_3), \quad (4.102a)$$

$$\|\mathbf{UH} - \mathbf{MU}^*(\mathbf{\Lambda}^*)^{-1}\|_{2,\infty} \leq 2(\mathcal{E}_{1,1} + \mathcal{E}_2 + \rho_1\mathcal{E}_3), \quad (4.102b)$$

$$\|\mathbf{U}\text{sgn}(\mathbf{H}) - \mathbf{U}^*\|_{2,\infty} \leq 4(\mathcal{E}_{1,1} + \mathcal{E}_2 + \mathcal{E}_3) + \frac{4c_2^2\sigma^2\sqrt{\mu rn}}{(\lambda_r^*)^2}, \quad (4.102c)$$

$$\begin{aligned} & \|\mathbf{U}\text{sgn}(\mathbf{H}) - \mathbf{MU}^*(\mathbf{\Lambda}^*)^{-1}\|_{2,\infty} \\ & \leq 3\mathcal{E}_{1,1} + 3\mathcal{E}_2 + \left(2\rho_1 + \frac{8c_2^2\sigma^2n}{(\lambda_r^*)^2}\right)\mathcal{E}_3 + \frac{4c_2^2\sigma^2\sqrt{\mu rn}}{(\lambda_r^*)^2}. \end{aligned} \quad (4.102d)$$

Remark 4.6. When \mathcal{E}_3 is the dominant term, the bound (4.102b) might be stronger than (4.102a) if ρ_1 is small.

With this lemma in mind, everything boils down to (i) establishing the relation (4.101) and (ii) deriving a tight bound on $\mathcal{E}_{1,1}$, which form the main content of the rest of the proof. In light of the triangle inequality

$$\|\mathbf{M}(\mathbf{UH} - \mathbf{U}^*)\|_{2,\infty} \leq \|\mathbf{E}(\mathbf{UH} - \mathbf{U}^*)\|_{2,\infty} + \|\mathbf{M}^*(\mathbf{UH} - \mathbf{U}^*)\|_{2,\infty},$$

we dedicate the next two steps to bounding $\|\mathbf{E}(\mathbf{UH} - \mathbf{U}^*)\|_{2,\infty}$ and $\|\mathbf{M}^*(\mathbf{UH} - \mathbf{U}^*)\|_{2,\infty}$ respectively.

Step 2: bounding $\|\mathbf{E}(\mathbf{UH} - \mathbf{U}^*)\|_{2,\infty}$ via leave-one-out analysis

To obtain tight row-wise control of $\mathbf{E}(\mathbf{UH} - \mathbf{U}^*)$, one needs to carefully decouple the statistical dependency between \mathbf{E} and \mathbf{U} , which is where the leave-one-out idea comes into play.

Step 2.1: a convenient decomposition. We start by invoking the triangle inequality to decompose the target quantity as follows

$$\begin{aligned} \|\mathbf{E}(\mathbf{UH} - \mathbf{U}^*)\|_{2,\infty} &= \max_l \|\mathbf{E}_{l,\cdot}(\mathbf{UH} - \mathbf{U}^*)\|_2 \\ &\leq \max_l \left\{ \|\mathbf{E}_{l,\cdot}(\mathbf{U}^{(l)}\mathbf{H}^{(l)} - \mathbf{U}^*)\|_2 + \|\mathbf{E}_{l,\cdot}(\mathbf{UH} - \mathbf{U}^{(l)}\mathbf{H}^{(l)})\|_2 \right\} \\ &\leq \max_l \left\{ \|\mathbf{E}_{l,\cdot}(\mathbf{U}^{(l)}\mathbf{H}^{(l)} - \mathbf{U}^*)\|_2 + \|\mathbf{E}\|\|\mathbf{UH} - \mathbf{U}^{(l)}\mathbf{H}^{(l)}\|_{\text{F}} \right\}. \end{aligned} \tag{4.103}$$

In words, when controlling the l -th row of $\mathbf{E}(\mathbf{UH} - \mathbf{U}^*)$, we attempt to employ $\mathbf{U}^{(l)}\mathbf{H}^{(l)}$ as a surrogate of \mathbf{UH} . The benefits to be harvested from this decomposition are:

- The statistical independence between $\mathbf{E}_{l,\cdot}$ and $\mathbf{U}^{(l)}\mathbf{H}^{(l)}$ allows for convenient upper bounds on $\|\mathbf{E}_{l,\cdot}(\mathbf{U}^{(l)}\mathbf{H}^{(l)} - \mathbf{U}^*)\|_2$;
- \mathbf{UH} and $\mathbf{U}^{(l)}\mathbf{H}^{(l)}$ are expected to be exceedingly close, so that the discrepancy incurred by replacing \mathbf{UH} with $\mathbf{U}^{(l)}\mathbf{H}^{(l)}$ is negligible.

In what follows, we flesh out the proof details.

Step 2.2: the proximity of \mathbf{UH} and $\mathbf{U}^{(l)}\mathbf{H}^{(l)}$. Given that

$$\begin{aligned}\|\mathbf{UH} - \mathbf{U}^{(l)}\mathbf{H}^{(l)}\|_{\text{F}} &= \|\mathbf{UU}^{\top}\mathbf{U}^* - \mathbf{U}^{(l)}\mathbf{U}^{(l)\top}\mathbf{U}^*\|_{\text{F}} \\ &\leq \|\mathbf{UU}^{\top} - \mathbf{U}^{(l)}\mathbf{U}^{(l)\top}\|_{\text{F}}\|\mathbf{U}^*\| \\ &= \|\mathbf{UU}^{\top} - \mathbf{U}^{(l)}\mathbf{U}^{(l)\top}\|_{\text{F}},\end{aligned}\quad (4.104)$$

it boils down to bounding $\|\mathbf{UU}^{\top} - \mathbf{U}^{(l)}\mathbf{U}^{(l)\top}\|_{\text{F}}$. Under simple conditions on the eigengap and the perturbation size (see (4.95) in Remark 4.4), the Davis-Kahan theorem (cf. Corollary 2.8) yields

$$\begin{aligned}\|\mathbf{UU}^{\top} - \mathbf{U}^{(l)}\mathbf{U}^{(l)\top}\|_{\text{F}} &\leq \frac{2\|(\mathbf{M} - \mathbf{M}^{(l)})\mathbf{U}^{(l)}\|_{\text{F}}}{|\lambda_r^{(l)}| - |\lambda_{r+1}^{(l)}|} \\ &\leq \frac{4\|(\mathbf{M} - \mathbf{M}^{(l)})\mathbf{U}^{(l)}\|_{\text{F}}}{\lambda_r^*}.\end{aligned}\quad (4.105)$$

It remains to develop an upper bound on $\|(\mathbf{M} - \mathbf{M}^{(l)})\mathbf{U}^{(l)}\|$. The way we construct $\mathbf{M}^{(l)}$ (see Section 4.8.1) allows us to express

$$(\mathbf{M} - \mathbf{M}^{(l)})\mathbf{U}^{(l)} = \mathbf{e}_l \mathbf{E}_{l,\cdot} \mathbf{U}^{(l)} + (\mathbf{E}_{\cdot,l} - E_{l,l} \mathbf{e}_l) \mathbf{e}_l^{\top} \mathbf{U}^{(l)}.$$

This together with the triangle inequality and the fact (4.97) gives

$$\begin{aligned}\|(\mathbf{M} - \mathbf{M}^{(l)})\mathbf{U}^{(l)}\|_{\text{F}} &\leq \|\mathbf{E}_{l,\cdot} \mathbf{U}^{(l)}\|_2 + \|\mathbf{E}_{\cdot,l} - E_{l,l} \mathbf{e}_l\|_2 \|\mathbf{U}^{(l)}\|_{2,\infty} \\ &\leq \|\mathbf{E}_{l,\cdot} \mathbf{U}^{(l)}\|_2 + 2\|\mathbf{E}\| \|\mathbf{U}^{(l)}\mathbf{H}^{(l)}\|_{2,\infty} \\ &\leq \|\mathbf{E}_{l,\cdot} \mathbf{U}^{(l)}\|_2 + 2\|\mathbf{E}\| \|\mathbf{UH}\|_{2,\infty} + 2\|\mathbf{E}\| \|\mathbf{UH} - \mathbf{U}^{(l)}\mathbf{H}^{(l)}\|_{\text{F}}.\end{aligned}$$

Substitution into (4.104) and (4.105) gives

$$\begin{aligned}\|\mathbf{UH} - \mathbf{U}^{(l)}\mathbf{H}^{(l)}\|_{\text{F}} &\leq \|\mathbf{UU}^{\top} - \mathbf{U}^{(l)}\mathbf{U}^{(l)\top}\|_{\text{F}} \\ &\leq \frac{4\|\mathbf{E}_{l,\cdot} \mathbf{U}^{(l)}\|_2 + 8\|\mathbf{E}\| \|\mathbf{UH}\|_{2,\infty} + 8\|\mathbf{E}\| \|\mathbf{UH} - \mathbf{U}^{(l)}\mathbf{H}^{(l)}\|_{\text{F}}}{\lambda_r^*}.\end{aligned}$$

As long as $\|\mathbf{E}\|/\lambda_r^* \leq 1/16$, one can further rearrange terms to obtain

$$\|\mathbf{UH} - \mathbf{U}^{(l)}\mathbf{H}^{(l)}\|_{\text{F}} \leq \frac{8\|\mathbf{E}_{l,\cdot} \mathbf{U}^{(l)}\|_2 + 16\|\mathbf{E}\| \|\mathbf{UH}\|_{2,\infty}}{\lambda_r^*}.\quad (4.106)$$

In addition, the fact (4.97) combined with the triangle inequality yields

$$\frac{1}{2}\|\mathbf{E}_{l,\cdot} \mathbf{U}^{(l)}\|_2 \leq \|\mathbf{E}_{l,\cdot} \mathbf{U}^{(l)}\mathbf{H}^{(l)}\|_2 \leq \|\mathbf{E}_{l,\cdot}(\mathbf{U}^{(l)}\mathbf{H}^{(l)} - \mathbf{U}^*)\|_2 + \|\mathbf{E}_{l,\cdot} \mathbf{U}^*\|_2,$$

which taken collectively with (4.106) reveals that

$$\begin{aligned} \|\mathbf{UH} - \mathbf{U}^{(l)}\mathbf{H}^{(l)}\|_{\text{F}} &\leq \frac{16\|\mathbf{E}_{l,\cdot}(\mathbf{U}^{(l)}\mathbf{H}^{(l)} - \mathbf{U}^*)\|_2}{\lambda_r^*} \\ &+ \frac{16\left\{\|\mathbf{E}_{l,\cdot}\mathbf{U}^*\|_2 + \|\mathbf{E}\|\|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty} + \|\mathbf{E}\|\|\mathbf{U}^*\|_{2,\infty}\right\}}{\lambda_r^*}. \end{aligned} \quad (4.107)$$

Step 2.3: bounding $\|\mathbf{E}_{l,\cdot}(\mathbf{U}^{(l)}\mathbf{H}^{(l)} - \mathbf{U}^*)\|_2$. Recognizing that $\mathbf{E}_{l,\cdot}$ is statistically independent of $\mathbf{U}^{(l)}$ (since $\mathbf{U}^{(l)}$ is computed without using $\mathbf{E}_{l,\cdot}$), we invoke Lemma 4.13 (more precisely, we use the proof of this lemma) to demonstrate that with probability exceeding $1 - 2n^{-6}$,

$$\begin{aligned} &\|\mathbf{E}_{l,\cdot}(\mathbf{U}^{(l)}\mathbf{H}^{(l)} - \mathbf{U}^*)\|_2 \\ &\leq 4\sigma\sqrt{\log n}\|\mathbf{U}^{(l)}\mathbf{H}^{(l)} - \mathbf{U}^*\|_{\text{F}} + (6B\log n)\|\mathbf{U}^{(l)}\mathbf{H}^{(l)} - \mathbf{U}^*\|_{2,\infty} \\ &\leq 4\sigma\sqrt{\log n}\|\mathbf{UH} - \mathbf{U}^*\|_{\text{F}} + (6B\log n)\|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty} \\ &\quad + (10B\log n)\|\mathbf{UH} - \mathbf{U}^{(l)}\mathbf{H}^{(l)}\|_{\text{F}} \end{aligned} \quad (4.108)$$

holds simultaneously for all $1 \leq l \leq n$, where the last line results from the triangle inequality and the fact $4\sigma\sqrt{\log n} + 6B\log n \leq 10B\log n$.

Step 2.4: combining the above bounds. The careful reader would immediately remark that the inequalities (4.107) and (4.108) are convoluted, both of which involve the terms $\|\mathbf{E}_{l,\cdot}(\mathbf{U}^{(l)}\mathbf{H}^{(l)} - \mathbf{U}^*)\|_2$ and $\|\mathbf{UH} - \mathbf{U}^{(l)}\mathbf{H}^{(l)}\|_{\text{F}}$. Fortunately, one can substitute (4.108) into (4.107) to produce a cleaner bound. By doing so and exploiting the condition $320B\log n \leq \lambda_r^*$, we rearrange terms to reach

$$\begin{aligned} \|\mathbf{UH} - \mathbf{U}^{(l)}\mathbf{H}^{(l)}\|_{\text{F}} &\leq \frac{32\|\mathbf{E}_{l,\cdot}\mathbf{U}^*\|_2 + 32\|\mathbf{E}\|\|\mathbf{U}^*\|_{2,\infty}}{\lambda_r^*} \\ &+ \frac{128\sigma\sqrt{\log n}\|\mathbf{UH} - \mathbf{U}^*\|_{\text{F}} + (32c_2\sigma\sqrt{n} + 192B\log n)\|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty}}{\lambda_r^*}, \end{aligned}$$

where we have also used the upper bound on $\|\mathbf{E}\|$ derived in (4.94). Meanwhile, plugging the above inequality into (4.108) yields

$$\|\mathbf{E}_{l,\cdot}(\mathbf{U}^{(l)}\mathbf{H}^{(l)} - \mathbf{U}^*)\|_2 \leq \frac{320B\log n}{\lambda_r^*}(\|\mathbf{E}_{l,\cdot}\mathbf{U}^*\|_2 + \|\mathbf{E}\|\|\mathbf{U}^*\|_{2,\infty})$$

$$+ 5\sigma\sqrt{\log n} \|\mathbf{UH} - \mathbf{U}^*\|_{\text{F}} + (7B\log n) \|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty}, \quad (4.109)$$

provided that $\max\{\sigma\sqrt{n}, B\log n\} \leq c_3\lambda_r^*$ for some small constant c_3 .

Substituting the preceding two bounds into (4.103) and combining terms reveal the existence of some constant $c_4 > 0$ such that

$$\begin{aligned} \|\mathbf{E}(\mathbf{UH} - \mathbf{U}^*)\|_{2,\infty} \\ \leq \alpha_0 + \alpha_1 + c_4(\sigma\sqrt{n} + B\log n) \|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty} \end{aligned} \quad (4.110)$$

provided that $\max\{\sigma\sqrt{n}, B\log n\} \leq c_3\lambda_r^*$ for some constant $c_3 > 0$ small enough, where

$$\begin{aligned} \alpha_0 &:= \frac{32c_2\sigma\sqrt{n} + 320B\log n}{\lambda_r^*} (\|\mathbf{EU}^*\|_{2,\infty} + \|\mathbf{E}\| \|\mathbf{U}^*\|_{2,\infty}), \\ \alpha_1 &:= 6\sigma\sqrt{\log n} \|\mathbf{UH} - \mathbf{U}^*\|_{\text{F}}. \end{aligned} \quad (4.111)$$

Before continuing, note that we are already well-equipped to bound the above two quantities. First, α_0 can be bounded by

$$\begin{aligned} \alpha_0 &\leq \frac{32c_2\sigma\sqrt{n} + 320B\log n}{\lambda_r^*} \\ &\quad \cdot \left(4\sigma\sqrt{r\log n} + 6B\sqrt{\frac{\mu r}{n}\log n} + c_2\sigma\sqrt{\mu r} \right), \end{aligned} \quad (4.112)$$

where we have used the bounds concerning \mathbf{E} from Lemma 4.13 as well as the definition (4.24). Regarding α_1 , it is seen from (4.94e) that

$$\alpha_1 \leq \frac{12c_2\sigma^2\sqrt{rn\log n}}{\lambda_r^*}. \quad (4.113)$$

Step 3: bounding $\|\mathbf{M}^*(\mathbf{UH} - \mathbf{U}^*)\|_{2,\infty}$

We make the key observation that

$$\begin{aligned} \|\mathbf{M}^*(\mathbf{UH} - \mathbf{U}^*)\|_{2,\infty} &= \|\mathbf{U}^*\mathbf{\Lambda}^*\mathbf{U}^{*\top}(\mathbf{UH} - \mathbf{U}^*)\|_{2,\infty} \\ &\leq \|\mathbf{U}^*\|_{2,\infty} \|\mathbf{\Lambda}^*\| \|\mathbf{U}^{*\top}(\mathbf{UH} - \mathbf{U}^*)\| \\ &= \sqrt{\frac{\mu r}{n}} \|\mathbf{\Lambda}^*\| \|\mathbf{U}\mathbf{U}^\top - \mathbf{U}^*\mathbf{U}^{*\top}\|^2. \end{aligned} \quad (4.114)$$

Here, the last identity holds true due to the following observation

$$\|\mathbf{U}^{*\top}(\mathbf{UH} - \mathbf{U}^*)\| = \|\mathbf{U}^{*\top}\mathbf{U}\mathbf{U}^\top\mathbf{U}^* - \mathbf{U}^{*\top}\mathbf{U}^*\| = \|\mathbf{U}^{*\top}\mathbf{U}\mathbf{U}^\top\mathbf{U}^* - \mathbf{I}\|$$

$$\begin{aligned}
&= \|\mathbf{Y}(\cos^2 \Theta) \mathbf{Y}^\top - \mathbf{I}\| = \|\cos^2 \Theta - \mathbf{I}\| \\
&= \|\sin^2 \Theta\| = \|\sin \Theta\|^2,
\end{aligned}$$

where we denote by $\mathbf{X}(\cos \Theta) \mathbf{Y}^\top$ the SVD of $\mathbf{U}^\top \mathbf{U}^*$, with \mathbf{X} and \mathbf{Y} being orthonormal matrices and Θ the diagonal matrix consisting of the principal angles between \mathbf{U} and \mathbf{U}^* (see the definition in (2.5)). The above bounds combined with (4.94b) indicate that

$$\begin{aligned}
\|\mathbf{M}^*(\mathbf{UH} - \mathbf{U}^*)\|_{2,\infty} &\leq |\lambda_1^*| \sqrt{\frac{\mu r}{n}} \|\sin \Theta\|^2 \\
&\leq \frac{4c_2^2 \kappa \sigma^2 \sqrt{\mu r n}}{\lambda_r^*} =: \alpha_2.
\end{aligned} \tag{4.115}$$

Step 4: putting all pieces together

Combining the bounds (4.110) and (4.115) in Steps 2-3 and using the definition of \mathcal{E}_1 (see Lemma 4.16) give

$$\begin{aligned}
\mathcal{E}_1 &\leq \frac{2\|\mathbf{E}(\mathbf{UH} - \mathbf{U}^*)\|_{2,\infty} + 2\|\mathbf{M}^*(\mathbf{UH} - \mathbf{U}^*)\|_{2,\infty}}{\lambda_r^*} \\
&\leq \mathcal{E}_{1,1} + \rho_1 \|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty},
\end{aligned} \tag{4.116}$$

where

$$\mathcal{E}_{1,1} := \frac{2(\alpha_0 + \alpha_1 + \alpha_2)}{\lambda_r^*} \quad \text{and} \quad \rho_1 := \frac{2c_4(\sigma\sqrt{n} + B \log n)}{\lambda_r^*}. \tag{4.117}$$

This matches precisely the relation hypothesized in (4.101). In particular, one has $0 < \rho \leq 1/2$ as long as $4c_4(\sigma\sqrt{n} + B \log n) \leq \lambda_r^*$, which holds whenever $\sigma\sqrt{n} \log n \leq c_\sigma \lambda_r^*$ for a sufficiently small $c_\sigma > 0$ in view of our assumption on B (cf. (4.29)).

Recall that $\mathcal{E}_{1,1}$, \mathcal{E}_2 , \mathcal{E}_3 , α_0 , α_1 , α_2 and ρ_1 have been controlled in (4.117), (4.99), (4.100), (4.112), (4.113), (4.115) and (4.117), respectively. In addition, recall our assumption (4.29) and suppose that $\sigma\sqrt{n} \leq c_\sigma \lambda_r^*$ for some sufficiently small constant $c_\sigma > 0$. With these bounds and assumptions in mind, invoking Lemma 4.17 and combining terms immediately conclude the proof.

Remark 4.7. We shall also make note of an immediate consequence of the above argument as follows

$$\|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty} \lesssim \frac{\sigma\kappa\sqrt{\mu r} + \sigma\sqrt{r\log n}}{\lambda_r^*}, \quad (4.118)$$

which will prove useful for deriving other important results.

4.8.4 Proof of auxiliary lemmas

Proof of Lemma 4.13. Under Assumption 4.1, Theorem 3.4 (in particular (3.9)) reveals the existence of some constant $c_2 > 0$ such that

$$\max_l \|\mathbf{E}^{(l)}\| \leq \|\mathbf{E}\| \leq c_2\sigma\sqrt{n},$$

holds with probability exceeding $1 - O(n^{-7})$.

When it comes to \mathbf{EA} , we proceed by viewing its l -th row $\mathbf{E}_{l,\cdot}\mathbf{A}$ as a sum of independent random vectors as follows

$$\mathbf{E}_{l,\cdot}\mathbf{A} = \sum_j E_{l,j} \mathbf{A}_{j,\cdot} =: \sum_j \mathbf{z}_j,$$

which can be controlled by the matrix Bernstein inequality. Specifically, it is seen from Assumption 4.1 that

$$\begin{aligned} v &:= \sum_j \mathbb{E}[\|\mathbf{z}_j\|_2^2] \leq \sigma^2 \sum_j \|\mathbf{A}_{j,\cdot}\|_2^2 = \sigma^2 \|\mathbf{A}\|_{\text{F}}^2; \\ L &:= \max_j \|\mathbf{z}_j\|_2 = \max_j |E_{l,j}| \|\mathbf{Z}_{j,\cdot}\|_2 \leq B \|\mathbf{Z}\|_{2,\infty}. \end{aligned}$$

Invoke the matrix Bernstein inequality (cf. Corollary 3.3) and take the union bound to demonstrate that: with probability exceeding $1 - 2n^{-6}$,

$\|\mathbf{E}_{l,\cdot}\mathbf{A}\|_2 \leq 4\sqrt{v\log n} + 6L\log n \leq 4\sigma\sqrt{\log n} \|\mathbf{A}\|_{\text{F}} + (6B\log n) \|\mathbf{A}\|_{2,\infty}$ holds simultaneously for all $1 \leq l \leq n$, thus concluding the proof.

Proof of Lemma 4.14. Lemma 4.13 tells us that

$$\max_l \|\mathbf{E}^{(l)}\| \leq \|\mathbf{E}\| \leq c_2\sigma\sqrt{n} \quad (4.119)$$

holds with probability exceeding $1 - O(n^{-7})$. Repeating the argument in the proof of Corollary 2.8 reveals that

$$\max_{1 \leq j \leq r} |\lambda_j| \geq \lambda_r^* - \|\mathbf{E}\|, \quad \max_{1 \leq j \leq r} |\lambda_j^{(l)}| \geq \lambda_r^* - \|\mathbf{E}^{(l)}\|, \quad (4.120a)$$

$$\max_{j:j>r} |\lambda_j| \leq \|\mathbf{E}\|, \quad \max_{j:j>r} |\lambda_j^{(l)}| \leq \|\mathbf{E}^{(l)}\|, \quad (4.120b)$$

which taken together with (4.119) validates (4.94c)-(4.94d).

Regarding (4.94a) and (4.94b), apply Corollary 2.8 to reach

$$\text{dist}(\mathbf{U}, \mathbf{U}^*) \leq \sqrt{2} \|\sin \Theta\| \leq \frac{2\|\mathbf{E}\|}{\lambda_r^*} \leq \frac{2c_2\sigma\sqrt{n}}{\lambda_r^*} \quad (4.121)$$

as long as $\|\mathbf{E}\| \leq c_2\sigma\sqrt{n} \leq (1 - 1/\sqrt{2})\lambda_r^*$, where we have used $\lambda_{r+1}^* = 0$. The bound on $\text{dist}(\mathbf{U}^{(l)}, \mathbf{U}^*)$ follows from the same argument.

Additionally, regarding $\mathbf{UH} - \mathbf{U}^*$ we can derive

$$\begin{aligned} \|\mathbf{UH} - \mathbf{U}^*\|_F &= \|\mathbf{U}\mathbf{U}^\top \mathbf{U}^* - \mathbf{U}^* \mathbf{U}^{*\top} \mathbf{U}^*\|_F \\ &\leq \|\mathbf{U}\mathbf{U}^\top - \mathbf{U}^* \mathbf{U}^{*\top}\| \|\mathbf{U}^*\|_F \\ &= \|\sin \Theta\| \|\mathbf{U}^*\|_F \leq \frac{2c_2\sigma\sqrt{rn}}{\lambda_r^*}, \end{aligned} \quad (4.122)$$

where the last line arises from Lemma 2.5, (4.94b), and $\|\mathbf{U}^*\|_F = \sqrt{r}$.

Proof of Lemma 4.15. We shall only prove the result for \mathbf{H} ; the proof for $\mathbf{H}^{(l)}$ follows from identical arguments and is hence omitted.

From our discussion in Section 2.2.2, one can express the SVD of $\mathbf{H} = \mathbf{U}^\top \mathbf{U}^*$ as $\mathbf{H} = \mathbf{X}(\cos \Theta)\mathbf{Y}^\top$, where the columns of \mathbf{X} (resp. \mathbf{Y}) are the left (resp. right) singular vectors of \mathbf{H} , and Θ is a diagonal matrix composed of the principal angles between \mathbf{U} and \mathbf{U}^* . In light of this and the definition (4.30), we can establish (4.96b) as follows

$$\begin{aligned} \|\mathbf{H} - \text{sgn}(\mathbf{H})\| &= \|\mathbf{X}(\cos \Theta - \mathbf{I})\mathbf{Y}^\top\| = \|\mathbf{I} - \cos \Theta\| \\ &\leq \|\mathbf{I} - \cos^2 \Theta\| = \|\sin \Theta\|^2 \\ &\leq \frac{2c_2^2\sigma^2 n}{(\lambda_r^*)^2}, \end{aligned} \quad (4.123)$$

where the middle line holds since $1 - \cos \theta \leq 1 - \cos^2 \theta$, and the last line follows from (4.94b).

Coming back to the claim (4.96a), it suffices to justify that $\sigma_{\min}(\mathbf{H}) \geq 1/2$. Recognizing that $\text{sgn}(\mathbf{H}) = \mathbf{X}\mathbf{Y}^\top$, we see that all singular values of $\text{sgn}(\mathbf{H})$ equal 1. Thus, Weyl's inequality together with (4.123) gives

$$\sigma_{\min}(\mathbf{H}) \geq \sigma_{\min}(\text{sgn}(\mathbf{H})) - \|\mathbf{H} - \text{sgn}(\mathbf{H})\| \geq 1 - \frac{2c_2^2\sigma^2 n}{(\lambda_r^*)^2} \geq \frac{1}{2},$$

with the proviso that $2c_2\sigma\sqrt{n} \leq \lambda_r^*$.

Proof of Lemma 4.16. We start by connecting $\mathbf{U}\mathbf{H}\Lambda^*$ more explicitly with $\mathbf{M}\mathbf{U}^*$ as follows (the invertibility of Λ can be deduced from (4.94c))

$$\mathbf{U}\mathbf{H}\Lambda^* = \mathbf{U}\Lambda\Lambda^{-1}\mathbf{H}\Lambda^* = \mathbf{M}\mathbf{U}\Lambda^{-1}\mathbf{U}^\top\mathbf{U}^*\Lambda^*, \quad (4.124)$$

which relies on the definition (4.89) and the eigendecomposition $\mathbf{M}\mathbf{U} = \mathbf{U}\Lambda$. In addition, the eigendecomposition $\mathbf{M}^*\mathbf{U}^* = \mathbf{U}^*\Lambda^*$ gives

$$\begin{aligned} \mathbf{U}^\top\mathbf{U}^*\Lambda^* &= \mathbf{U}^\top\mathbf{M}^*\mathbf{U}^* = \mathbf{U}^\top\mathbf{M}\mathbf{U}^* - \mathbf{U}^\top\mathbf{E}\mathbf{U}^* \\ &= \mathbf{U}\Lambda^\top\mathbf{U}^* - \mathbf{U}^\top\mathbf{E}\mathbf{U}^*, \end{aligned} \quad (4.125)$$

which taken collectively with (4.124) demonstrates that

$$\begin{aligned} \mathbf{U}\mathbf{H}\Lambda^* &= \mathbf{M}\mathbf{U}\Lambda^{-1}\Lambda\mathbf{U}^\top\mathbf{U}^* - \mathbf{M}\mathbf{U}\Lambda^{-1}\mathbf{U}^\top\mathbf{E}\mathbf{U}^* \\ &= \mathbf{M}\mathbf{U}\mathbf{H} - \mathbf{M}\mathbf{U}\Lambda^{-1}\mathbf{U}^\top\mathbf{E}\mathbf{U}^* \\ &= \mathbf{M}\mathbf{U}^* + \mathbf{M}(\mathbf{U}\mathbf{H} - \mathbf{U}^*) - \mathbf{M}\mathbf{U}\Lambda^{-1}\mathbf{U}^\top\mathbf{E}\mathbf{U}^*. \end{aligned}$$

Consequently, the difference between $\mathbf{U}\mathbf{H}\Lambda^*$ and $\mathbf{M}\mathbf{U}^*$ obeys

$$\begin{aligned} \|\mathbf{U}\mathbf{H}\Lambda^* - \mathbf{M}\mathbf{U}^*\|_{2,\infty} &\leq \|\mathbf{M}(\mathbf{U}\mathbf{H} - \mathbf{U}^*)\|_{2,\infty} \\ &\quad + \|\mathbf{M}\mathbf{U}\Lambda^{-1}\mathbf{U}^\top\mathbf{E}\mathbf{U}^*\|_{2,\infty}. \end{aligned} \quad (4.126)$$

Regarding the second term in (4.126), one can deduce that

$$\begin{aligned} &\|\mathbf{M}\mathbf{U}\Lambda^{-1}\mathbf{U}^\top\mathbf{E}\mathbf{U}^*\|_{2,\infty} \\ &\leq \|\mathbf{M}\mathbf{U}\|_{2,\infty} \|\Lambda^{-1}\| \cdot \|\mathbf{U}\| \cdot \|\mathbf{E}\| \cdot \|\mathbf{U}^*\| \\ &\stackrel{(i)}{\leq} \frac{2\|\mathbf{M}\mathbf{U}\|_{2,\infty} \|\mathbf{E}\|}{\lambda_r^*} \stackrel{(ii)}{\leq} \frac{4\|\mathbf{M}\mathbf{U}\mathbf{H}\|_{2,\infty} \|\mathbf{E}\|}{\lambda_r^*} \\ &\stackrel{(iii)}{\leq} \frac{4\|\mathbf{M}(\mathbf{U}\mathbf{H} - \mathbf{U}^*)\|_{2,\infty} \|\mathbf{E}\|}{\lambda_r^*} + \frac{4\|\mathbf{M}\mathbf{U}^*\|_{2,\infty} \|\mathbf{E}\|}{\lambda_r^*} \\ &\stackrel{(iv)}{\leq} \|\mathbf{M}(\mathbf{U}\mathbf{H} - \mathbf{U}^*)\|_{2,\infty} + \frac{4\|\mathbf{M}\mathbf{U}^*\|_{2,\infty} \|\mathbf{E}\|}{\lambda_r^*}. \end{aligned} \quad (4.127)$$

Here, (i) follows from the facts $\|\mathbf{U}\| = \|\mathbf{U}^*\| = 1$ and $|\lambda_r| \geq \lambda_r^* - c_2\sigma\sqrt{n} \geq \lambda_r^*/2$ (see Lemma 4.14), (ii) holds due to (4.97), (iii) invokes

the triangle inequality, whereas (iv) holds true provided that $4\|\mathbf{E}\| \leq \lambda_r^*$. Combine (4.126) and (4.127) to reach

$$\|\mathbf{UH}\Lambda^* - \mathbf{MU}^*\|_{2,\infty} \leq 2\|\mathbf{M}(\mathbf{UH} - \mathbf{U}^*)\|_{2,\infty} + \frac{4\|\mathbf{MU}^*\|_{2,\infty}\|\mathbf{E}\|}{\lambda_r^*},$$

which together with the fact that $\|(\Lambda^*)^{-1}\| = 1/\lambda_r^*$ and the elementary relation $\|\mathbf{A}\|_{2,\infty} = \|\mathbf{A}\Lambda^*(\Lambda^*)^{-1}\|_{2,\infty} \leq \|\mathbf{A}\Lambda^*\|_{2,\infty}\|(\Lambda^*)^{-1}\|$ yields the desired claim (4.98a).

When it comes to the second claim (4.98b), combining (4.98a) with the triangle inequality

$$\begin{aligned} \|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty} &= \|\mathbf{UH} - \mathbf{M}^*\mathbf{U}^*(\Lambda^*)^{-1}\|_{2,\infty} \\ &\leq \|\mathbf{UH} - \mathbf{MU}^*(\Lambda^*)^{-1}\|_{2,\infty} + \|\mathbf{EU}^*(\Lambda^*)^{-1}\|_{2,\infty} \\ &\leq \|\mathbf{UH} - \mathbf{MU}^*(\Lambda^*)^{-1}\|_{2,\infty} + \|\mathbf{EU}^*\|_{2,\infty} / \lambda_r^* \end{aligned}$$

immediately establishes the advertised bound. Here the last relation again arises from the elementary inequality $\|\mathbf{AB}\|_{2,\infty} \leq \|\mathbf{A}\|_{2,\infty}\|\mathbf{B}\|$.

Proof of Lemma 4.17. First of all, taking Condition (4.101) collectively with (4.98b) and rearranging terms yield (4.102a):

$$\|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty} \leq \frac{1}{1 - \rho_1}(\mathcal{E}_{1,1} + \mathcal{E}_2 + \mathcal{E}_3) \leq 2(\mathcal{E}_{1,1} + \mathcal{E}_2 + \mathcal{E}_3), \quad (4.128)$$

where the last inequality follows from $\rho \leq 1/2$. Substituting (4.101) and (4.102a) into (4.98a) then gives (4.102b):

$$\begin{aligned} \|\mathbf{UH} - \mathbf{MU}^*(\Lambda^*)^{-1}\|_{2,\infty} &\leq \mathcal{E}_{1,1} + 2\rho_1(\mathcal{E}_{1,1} + \mathcal{E}_2 + \mathcal{E}_3) + \mathcal{E}_2 \\ &\leq 2\mathcal{E}_{1,1} + 2\mathcal{E}_2 + 2\rho_1\mathcal{E}_3, \end{aligned} \quad (4.129)$$

where once again we use the assumption that $\rho \leq 1/2$. In addition, the following observation connects \mathbf{UH} with $\mathbf{U}\text{sgn}(\mathbf{H})$:

$$\begin{aligned} \|\mathbf{UH} - \mathbf{U}\text{sgn}(\mathbf{H})\|_{2,\infty} &\leq \|\mathbf{U}\|_{2,\infty}\|\mathbf{H} - \text{sgn}(\mathbf{H})\| \\ &\stackrel{(i)}{\leq} \frac{2c_2^2\sigma^2n}{(\lambda_r^*)^2}\|\mathbf{U}\|_{2,\infty} \stackrel{(ii)}{\leq} \frac{4c_2^2\sigma^2n}{(\lambda_r^*)^2}\|\mathbf{UH}\|_{2,\infty} \\ &\stackrel{(iii)}{\leq} \frac{4c_2^2\sigma^2n}{(\lambda_r^*)^2}\|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty} + \frac{4c_2^2\sigma^2n}{(\lambda_r^*)^2}\sqrt{\frac{\mu r}{n}}, \end{aligned} \quad (4.130)$$

where (i) results from (4.96b), (ii) relies on (4.97a), and (iii) comes from the triangle inequality $\|\mathbf{UH}\|_{2,\infty} \leq \|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty} + \|\mathbf{U}^*\|_{2,\infty}$ and the definition (4.24).

The preceding bound together with the triangle inequality gives (4.102c):

$$\begin{aligned} \|\mathbf{U}\text{sgn}(\mathbf{H}) - \mathbf{U}^*\|_{2,\infty} &\leq \|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty} + \|\mathbf{UH} - \mathbf{U}\text{sgn}(\mathbf{H})\|_{2,\infty} \\ &\leq \left(1 + \frac{4c_2^2\sigma^2n}{(\lambda_r^*)^2}\right) \|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty} + \frac{4c_2^2\sigma^2n}{(\lambda_r^*)^2} \sqrt{\frac{\mu r}{n}} \\ &\leq 4(\mathcal{E}_{1,1} + \mathcal{E}_2 + \mathcal{E}_3) + \frac{4c_2^2\sigma^2\sqrt{\mu rn}}{(\lambda_r^*)^2}, \end{aligned}$$

where the last inequality holds as long as $4c_2^2\sigma^2n \leq (\lambda_r^*)^2$. Additionally, the inequalities (4.129) and (4.130) further allow us to deduce (4.102d):

$$\begin{aligned} &\|\mathbf{U}\text{sgn}(\mathbf{H}) - \mathbf{MU}^*(\Lambda^*)^{-1}\|_{2,\infty} \\ &\leq \|\mathbf{UH} - \mathbf{MU}^*(\Lambda^*)^{-1}\|_{2,\infty} + \|\mathbf{UH} - \mathbf{U}\text{sgn}(\mathbf{H})\|_{2,\infty} \\ &\leq 2\mathcal{E}_{1,1} + 2\mathcal{E}_2 + 2\rho_1\mathcal{E}_3 + \frac{8c_2^2\sigma^2n}{(\lambda_r^*)^2}(\mathcal{E}_{1,1} + \mathcal{E}_2 + \mathcal{E}_3) + \frac{4c_2^2\sigma^2\sqrt{\mu rn}}{(\lambda_r^*)^2} \\ &\leq 3\mathcal{E}_{1,1} + 3\mathcal{E}_2 + \left(2\rho_1 + \frac{8c_2^2\sigma^2n}{(\lambda_r^*)^2}\right)\mathcal{E}_3 + \frac{4c_2^2\sigma^2\sqrt{\mu rn}}{(\lambda_r^*)^2}. \end{aligned}$$

Here, the penultimate line combines (4.102a), (4.102b) and (4.130), while the last inequality relies on the assumption $8c_2^2\sigma^2n \leq (\lambda_r^*)^2$.

4.9 Appendix B: Proof of Corollary 4.3

Moving on to the proof of Corollary 4.3, we start by pointing out the main issue that deserves particular attention. Roughly speaking, we have learned from Theorem 4.2 (and its analysis) that $\mathbf{U}^* \approx \mathbf{UH}$ under mild conditions, which naturally suggests that

$$\mathbf{M}^* = \mathbf{U}^* \Lambda^* \mathbf{U}^{*\top} \approx \mathbf{UH} \Lambda^* \mathbf{H}^\top \mathbf{U}^\top.$$

As a result, in order to enable $\mathbf{M}^* \approx \mathbf{U}\Lambda\mathbf{U}^\top$, one would need to ensure $\mathbf{H}\Lambda^*\mathbf{H}^\top \approx \Lambda$.

The above argument, while highly informal, reveals the core idea underlying the proof. Our proof is based upon the following observation

$$\begin{aligned} \|\mathbf{U}\Lambda\mathbf{U}^\top - \mathbf{U}^*\Lambda^*\mathbf{U}^{*\top}\|_\infty &\leq \underbrace{\|\mathbf{U}(\Lambda - \mathbf{H}\Lambda^*\mathbf{H}^\top)\mathbf{U}^\top\|_\infty}_{=:\gamma_1} \\ &\quad + \underbrace{\|\mathbf{U}\mathbf{H}\Lambda^*\mathbf{H}^\top\mathbf{U}^\top - \mathbf{U}^*\Lambda^*\mathbf{U}^{*\top}\|_\infty}_{=:\gamma_2}, \end{aligned} \quad (4.131)$$

which leaves us with two terms to cope with.

Step 1: bounding γ_1

Regarding γ_1 defined in (4.131), it is seen that

$$\gamma_1 \leq \|\mathbf{U}\|_{2,\infty}^2 \|\Lambda - \mathbf{H}\Lambda^*\mathbf{H}^\top\| \leq \frac{4\mu r}{n} \|\Lambda - \mathbf{H}\Lambda^*\mathbf{H}^\top\|. \quad (4.132)$$

In the last relation, we have exploited the fact that

$$\begin{aligned} \|\mathbf{U}\|_{2,\infty} &= \|\mathbf{U}\text{sgn}(\mathbf{H})\|_{2,\infty} \leq \|\mathbf{U}^*\|_{2,\infty} + \|\mathbf{U}\text{sgn}(\mathbf{H}) - \mathbf{U}^*\|_{2,\infty} \\ &\leq 2\sqrt{\mu r/n}, \end{aligned} \quad (4.133)$$

where the last inequality relies on (4.31a) and the assumption $\sigma\sqrt{n}(\kappa + \sqrt{\log n}) \leq c_1\lambda_r^*$ for some sufficiently small constant $c_1 > 0$.

It then boils down to bounding $\|\Lambda - \mathbf{H}\Lambda^*\mathbf{H}^\top\|$. Towards this, it is seen from the identity (4.125) and the definition $\mathbf{H} = \mathbf{U}^\top\mathbf{U}^*$ that

$$\mathbf{H}\Lambda^*\mathbf{H}^\top - \Lambda = \Lambda\mathbf{H}\mathbf{H}^\top - \mathbf{U}^\top\mathbf{E}\mathbf{U}^*\mathbf{H}^\top - \Lambda,$$

which together with the triangle inequality reveals that

$$\|\mathbf{H}\Lambda^*\mathbf{H}^\top - \Lambda\| \leq \|\Lambda(\mathbf{H}\mathbf{H}^\top - \mathbf{I})\| + \|\mathbf{U}^\top\mathbf{E}\mathbf{U}^*\mathbf{H}^\top\|. \quad (4.134)$$

The rest of this step is devoted to controlling the above two terms.

With regards to the first term on the right-hand side of (4.134), we make the observation that

$$\|\mathbf{H}\mathbf{H}^\top - \mathbf{I}\| = \|\cos^2 \Theta - \mathbf{I}\| = \|\sin^2 \Theta\| \leq \frac{2c_2^2\sigma^2 n}{(\lambda_r^*)^2},$$

where, as usual, Θ denotes a diagonal matrix composed of the principal angles between \mathbf{U} and \mathbf{U}^* , and the last inequality results from

Lemma 4.14. This combined with Weyl's inequality and Lemma 4.14 leads to

$$\begin{aligned}\|\Lambda(\mathbf{H}\mathbf{H}^\top - \mathbf{I})\| &\leq \|\Lambda\| \|\mathbf{H}\mathbf{H}^\top - \mathbf{I}\| \leq (\|\Lambda^*\| + \|\mathbf{E}\|) \|\mathbf{H}\mathbf{H}^\top - \mathbf{I}\| \\ &\leq (|\lambda_1^*| + c_2\sigma\sqrt{n}) \frac{2c_2^2\sigma^2n}{(\lambda_r^*)^2} \leq \frac{4c_2^2\kappa\sigma^2n}{\lambda_r^*},\end{aligned}\quad (4.135)$$

provided that $c_2\sigma\sqrt{n} \leq \lambda_r^* \leq |\lambda_1^*|$.

When it comes to the second term on the right-hand side of (4.134), let us introduce an orthonormal matrix $\mathbf{R} := \arg \min_{Q \in \mathcal{O}^{r \times r}} \|\mathbf{U}Q - \mathbf{U}^*\|$, which helps us derive

$$\begin{aligned}\|\mathbf{U}^\top \mathbf{E} \mathbf{U}^* \mathbf{H}^\top\| &\leq \|\mathbf{U}^\top \mathbf{E} \mathbf{U}^*\| = \|\mathbf{R}^\top \mathbf{U}^\top \mathbf{E} \mathbf{U}^*\| \\ &\leq \|\mathbf{U}^{*\top} \mathbf{E} \mathbf{U}^*\| + \|(\mathbf{U}\mathbf{R} - \mathbf{U}^*)^\top \mathbf{E} \mathbf{U}^*\| \\ &\leq \|\mathbf{U}^{*\top} \mathbf{E} \mathbf{U}^*\| + \|\mathbf{E}\| \text{dist}(\mathbf{U}, \mathbf{U}^*).\end{aligned}\quad (4.136)$$

Here, the first inequality holds since $\|\mathbf{H}\| = \|\mathbf{U}^\top \mathbf{U}^*\| \leq 1$, while the last line follows since $\|\mathbf{U}^*\| = 1$ and $\|\mathbf{U}\mathbf{R} - \mathbf{U}^*\| = \text{dist}(\mathbf{U}, \mathbf{U}^*)$. In addition, we claim that with probability at least $1 - 2n^{-7}$,

$$\|\mathbf{U}^{*\top} \mathbf{E} \mathbf{U}^*\| \leq (6 + 12c_b)\sigma\sqrt{r \log n}.\quad (4.137)$$

If this claim were valid, then one could continue the derivation (4.136) and invoke Lemma 4.14 to demonstrate that

$$\|\mathbf{U}^\top \mathbf{E} \mathbf{U}^* \mathbf{H}^\top\| \leq (6 + 12c_b)\sigma\sqrt{r \log n} + \frac{2c_2^2\sigma^2n}{\lambda_r^*}.\quad (4.138)$$

To finish up, substituting (4.135) and (4.138) into (4.134) yields

$$\|\mathbf{H}\Lambda^* \mathbf{H}^\top - \Lambda\| \leq \frac{6c_2^2\kappa\sigma^2n}{\lambda_r^*} + (6 + 12c_b)\sigma\sqrt{r \log n},\quad (4.139)$$

which combined with (4.132) leads to

$$\gamma_1 \lesssim \frac{\sigma^2\kappa\mu r}{\lambda_r^*} + \frac{\sigma\mu\sqrt{r^3 \log n}}{n}.\quad (4.140)$$

Step 2: bounding γ_2

Before proceeding, we recall from (4.118) that

$$\|\mathbf{U}\mathbf{H} - \mathbf{U}^*\|_{2,\infty} \lesssim \frac{\sigma\kappa\sqrt{\mu r} + \sigma\sqrt{r \log n}}{\lambda_r^*} \lesssim \frac{\sigma\kappa\sqrt{\mu r \log n}}{\lambda_r^*}\quad (4.141)$$

Recognizing the basic decomposition

$$\gamma_2 = \|\mathbf{U} \mathbf{H} \boldsymbol{\Lambda}^* \mathbf{H}^\top \mathbf{U}^\top - \mathbf{U}^* \boldsymbol{\Lambda}^* \mathbf{U}^{*\top}\|_\infty = \|\mathbf{A}_1 + \mathbf{A}_1^\top + \mathbf{A}_2\|_\infty$$

with $\mathbf{A}_1 := (\mathbf{U} \mathbf{H} - \mathbf{U}^*) \boldsymbol{\Lambda}^* \mathbf{U}^{*\top}$ and $\mathbf{A}_2 := (\mathbf{U} \mathbf{H} - \mathbf{U}^*) \boldsymbol{\Lambda}^* (\mathbf{U} \mathbf{H} - \mathbf{U}^*)^\top$, we can control each of these terms separately. Firstly, observe that

$$\begin{aligned} \|\mathbf{A}_1\|_\infty &\leq \|\mathbf{U} \mathbf{H} - \mathbf{U}^*\|_{2,\infty} \|\mathbf{U}^*\|_{2,\infty} \|\boldsymbol{\Lambda}^*\| \\ &\lesssim |\lambda_1^*| \sqrt{\frac{\mu r}{n}} \cdot \frac{\sigma \kappa \sqrt{\mu r \log n}}{\lambda_r^*} \asymp \sigma \kappa^2 \mu r \sqrt{\frac{\log n}{n}}, \end{aligned}$$

where we have made use of (4.141). Similarly,

$$\begin{aligned} \|\mathbf{A}_2\|_\infty &\leq \|\mathbf{U} \mathbf{H} - \mathbf{U}^*\|_{2,\infty}^2 \|\boldsymbol{\Lambda}^*\| \lesssim |\lambda_1^*| \frac{\sigma^2 \kappa^2 \mu r \log n}{(\lambda_r^*)^2} \\ &\asymp \frac{\sigma^2 \kappa^3 \mu r \log n}{\lambda_r^*} \lesssim \sigma \kappa^2 \mu r \sqrt{\frac{\log n}{n}}, \end{aligned}$$

provided that $\sigma \kappa \sqrt{n \log n} \lesssim \lambda_r^*$. Consequently,

$$\gamma_2 \leq 2\|\mathbf{A}_1\|_\infty + \|\mathbf{A}_2\|_\infty \lesssim \sigma \kappa^2 \mu r \sqrt{\frac{\log n}{n}}.$$

Step 3: putting all this together

Combining the above bounds, we demonstrate that

$$\begin{aligned} \|\mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^\top - \mathbf{U}^* \boldsymbol{\Lambda}^* \mathbf{U}^{*\top}\|_\infty &\leq \gamma_1 + \gamma_2 \\ &\lesssim \sigma \kappa^2 \mu r \sqrt{\frac{\log n}{n}} + \frac{\kappa \mu r \sigma^2}{\lambda_r^*} + \frac{\sigma \mu r \sqrt{r \log n}}{n} \asymp \sigma \kappa^2 \mu r \sqrt{\frac{\log n}{n}}, \end{aligned} \quad (4.142)$$

provided that $\sigma \sqrt{n} \lesssim \lambda_r^*$. This concludes the proof of Corollary 4.3, as long as the claim (4.137) can be validated.

Proof of the claim (4.137)

Let us start by expressing $\mathbf{U}^{*\top} \mathbf{E} \mathbf{U}^*$ as a sum of independent random matrices as follows

$$\mathbf{U}^{*\top} \mathbf{E} \mathbf{U}^* = \sum_{i,j: i \geq j} E_{i,j} \left\{ (\mathbf{U}_{i,\cdot}^*)^\top \mathbf{U}_{j,\cdot}^* + (\mathbf{U}_{j,\cdot}^*)^\top \mathbf{U}_{i,\cdot}^* \right\} =: \sum_{i,j: i \geq j} \mathbf{Z}_{i,j}.$$

From the elementary inequality $(\mathbf{A} + \mathbf{A}^\top)^2 \preceq 2\mathbf{A}\mathbf{A}^\top + 2\mathbf{A}^\top\mathbf{A}$, we have

$$\begin{aligned}\mathbb{E}[\mathbf{Z}_{i,j}^2] &\preceq \sigma^2 \left\{ (\mathbf{U}_{i,\cdot}^*)^\top \mathbf{U}_{i,\cdot}^* + (\mathbf{U}_{j,\cdot}^*)^\top \mathbf{U}_{i,\cdot}^* \right\} \left\{ (\mathbf{U}_{i,\cdot}^*)^\top \mathbf{U}_{j,\cdot}^* + (\mathbf{U}_{j,\cdot}^*)^\top \mathbf{U}_{i,\cdot}^* \right\}^\top \\ &\preceq 2\sigma^2 \left\{ \|\mathbf{U}_{j,\cdot}^*\|_2^2 (\mathbf{U}_{i,\cdot}^*)^\top \mathbf{U}_{i,\cdot}^* + \|\mathbf{U}_{i,\cdot}^*\|_2^2 (\mathbf{U}_{j,\cdot}^*)^\top \mathbf{U}_{j,\cdot}^* \right\},\end{aligned}$$

thus indicating that

$$\begin{aligned}v &= \left\| \sum_{i,j: i \geq j} \mathbb{E}[\mathbf{Z}_{i,j}^2] \right\| \leq 2\sigma^2 \left\| \sum_{i=1}^n \sum_{j=1}^n \|\mathbf{U}_{j,\cdot}^*\|_2^2 (\mathbf{U}_{i,\cdot}^*)^\top \mathbf{U}_{i,\cdot}^* \right\| \\ &= 2\sigma^2 \|\mathbf{U}^*\|_{\text{F}}^2 \left\| \sum_i (\mathbf{U}_{i,\cdot}^*)^\top \mathbf{U}_{i,\cdot}^* \right\| = 2\sigma^2 r \|\mathbf{U}^{*\top} \mathbf{U}^*\| = 2\sigma^2 r.\end{aligned}$$

In addition, each matrix $\mathbf{Z}_{i,j}$ can be bounded in size by

$$\max_{i,j} \|\mathbf{Z}_{i,j}\| \leq 2 \max_{i,j} |E_{i,j}| \max_{i,j} \|\mathbf{U}_{i,\cdot}^*\|_2 \|\mathbf{U}_{j,\cdot}^*\|_2 \leq 2B \frac{\mu r}{n} =: L,$$

where we have used the definition of the incoherence parameter μ . Apply the matrix Bernstein inequality (see Corollary 3.3) to reach

$$\begin{aligned}\|\mathbf{U}^{*\top} \mathbf{E} \mathbf{U}^*\| &\leq 4\sqrt{v \log n} + 6L \log n \leq \sigma \sqrt{32r \log n} + 12B \frac{\mu r \log n}{n} \\ &\leq (6 + 12c_b)\sigma \sqrt{r \log n}\end{aligned}$$

with probability exceeding $1 - 2n^{-7}$. Here, the last line holds since

$$B \frac{\mu r \log n}{n} \leq c_b \sigma \sqrt{\frac{n}{\mu \log n}} \cdot \frac{\mu r \log n}{n} = c_b \sigma \sqrt{r \log n} \sqrt{\frac{\mu r}{n}} \leq c_b \sigma \sqrt{r \log n},$$

which relies on the assumption (4.29) and the basic fact $\mu \leq n/r$.

4.10 Appendix C: Proof of Theorem 4.4

A symmetrization trick. As alluded to previously, the proof is built on a “symmetric dilation” trick that helps symmetrize a general matrix. We start with the following definition.

Definition 4.3 (Symmetric dilation). For any matrix $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2}$, its symmetric dilation $\mathcal{S}(\mathbf{A}) \in \mathbb{R}^{(n_1+n_2) \times (n_1+n_2)}$ is defined to be

$$\mathcal{S}(\mathbf{A}) = \begin{bmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{A}^\top & \mathbf{0} \end{bmatrix}.$$

Apart from the symmetry of $\mathcal{S}(\mathbf{A})$, which is immediate from its definition, the main benefit of the symmetric dilation lies in the correspondence between the eigendecomposition of $\mathcal{S}(\mathbf{A})$ and the singular value decomposition of \mathbf{A} . More specifically, let $\mathbf{U}\Sigma\mathbf{V}^\top$ be the SVD of \mathbf{A} . Then one has the following eigendecomposition for $\mathcal{S}(\mathbf{A})$:

$$\mathcal{S}(\mathbf{A}) = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{U} & \mathbf{U} \\ \mathbf{V} & -\mathbf{V} \end{bmatrix} \cdot \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & -\Sigma \end{bmatrix} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{U} & \mathbf{U} \\ \mathbf{V} & -\mathbf{V} \end{bmatrix}^\top. \quad (4.143)$$

Here, the columns of $\frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{U} & \mathbf{U} \\ \mathbf{V} & -\mathbf{V} \end{bmatrix}$ are orthonormal and represent the eigenvectors of $\mathcal{S}(\mathbf{A})$, whereas $\begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & -\Sigma \end{bmatrix}$ contains all (non-zero) eigenvalues of $\mathcal{S}(\mathbf{A})$.

Utilizing this “symmetric dilation” trick, we can translate the observation model $\mathbf{M} = \mathbf{M}^* + \mathbf{E}$ into the following equivalent form

$$\mathcal{S}(\mathbf{M}) = \mathcal{S}(\mathbf{M}^*) + \mathcal{S}(\mathbf{E}),$$

which is in line with the symmetric observation model stated in (4.26).

Verifying conditions. To invoke the general theory in Section 4.2, one is required to first examine the spectral properties of $\mathcal{S}(\mathbf{M}^*)$ and $\mathcal{S}(\mathbf{M})$, as well as the assumptions on the noise part $\mathcal{S}(\mathbf{E})$.

Recall that $\mathbf{M}^* = \mathbf{U}^*\Sigma^*\mathbf{V}^{*\top}$, which together with the relation (4.143) reveals that: (i) $\mathcal{S}(\mathbf{M}^*)$ has rank $2r$ and condition number κ ; (ii) the nonzero eigenvalues of $\mathcal{S}(\mathbf{M}^*)$ and the corresponding eigenvectors are reflected respectively in the matrices

$$\bar{\Lambda}^* := \begin{bmatrix} \Sigma^* & \mathbf{0} \\ \mathbf{0} & -\Sigma^* \end{bmatrix} \quad \text{and} \quad \bar{\mathbf{U}}^* := \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{U}^* & \mathbf{U}^* \\ \mathbf{V}^* & -\mathbf{V}^* \end{bmatrix} \quad (4.144)$$

Similarly, given that the SVD of \mathbf{M} is $\mathbf{M} = \mathbf{U}\Sigma\mathbf{V}^\top + \mathbf{U}_\perp\Sigma_\perp\mathbf{V}_\perp^\top$, we see that the $2r$ -leading eigenvalues of $\mathcal{S}(\mathbf{M})$ and the corresponding eigenvectors are represented respectively by the matrices

$$\bar{\Lambda} := \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & -\Sigma \end{bmatrix} \quad \text{and} \quad \bar{\mathbf{U}} := \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{U} & \mathbf{U} \\ \mathbf{V} & -\mathbf{V} \end{bmatrix}.$$

Further, the incoherence parameter $\bar{\mu}$ of $\mathcal{S}(\mathbf{M}^*)$ (cf. (4.24)) obeys

$$\bar{\mu} := \frac{(n_1 + n_2)}{2r} \|\bar{\mathbf{U}}^*\|_{2,\infty}^2 \stackrel{(i)}{=} \frac{(n_1 + n_2)}{2r} \max \{ \|\mathbf{U}^*\|_{2,\infty}^2, \|\mathbf{V}^*\|_{2,\infty}^2 \}$$

$$\stackrel{\text{(ii)}}{\leq} \frac{(n_1 + n_2)}{\min\{n_1, n_2\}} \frac{\mu}{2} \stackrel{\text{(iii)}}{=} \frac{(n_1 + n_2)}{2n_1} \mu. \quad (4.145)$$

Here, the relation (i) is based on the definition (4.144), the inequality (ii) follows from the incoherence of \mathbf{M}^* (cf. Definition 3.1), while the last one (iii) holds under the assumption $n_1 \leq n_2$.

When it comes to the “symmetrized” noise part, it is straightforward to verify that under Assumption 4.2, the matrix $\mathcal{S}(\mathbf{E})$ satisfies Assumption 4.1 with precisely the quantities σ, B and c_b .

$\ell_{2,\infty}$ and ℓ_∞ guarantees. With the above preparations in place, apply Theorem 4.2 (more specifically (4.118)) to demonstrate that

$$\begin{aligned} \|\bar{\mathbf{U}}\bar{\mathbf{U}}^\top\bar{\mathbf{U}}^* - \bar{\mathbf{U}}^*\|_{2,\infty} &\lesssim \frac{\sigma\kappa\sqrt{\bar{\mu}r} + \sigma\sqrt{r\log n}}{\sigma_r^*} \\ &\lesssim \frac{\sigma\kappa\sqrt{n_2\mu r/n_1} + \sigma\sqrt{r\log n}}{\sigma_r^*}, \end{aligned} \quad (4.146)$$

where the last relation follows from (4.145). Further, note that

$$\bar{\mathbf{U}}\bar{\mathbf{U}}^\top\bar{\mathbf{U}}^* - \bar{\mathbf{U}}^* = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{U}\mathbf{H}_U - \mathbf{U}^*, & \mathbf{U}\mathbf{H}_U - \mathbf{U}^* \\ \mathbf{V}\mathbf{H}_V - \mathbf{V}^*, & -(\mathbf{V}\mathbf{H}_V - \mathbf{V}^*) \end{bmatrix},$$

where $\mathbf{H}_U := \mathbf{U}^\top\mathbf{U}^*$ and $\mathbf{H}_V := \mathbf{V}^\top\mathbf{V}^*$. Combining this with the upper bound (4.146) then yields

$$\max \{ \|\mathbf{U}\mathbf{H}_U - \mathbf{U}^*\|_{2,\infty}, \|\mathbf{V}\mathbf{H}_V - \mathbf{V}^*\|_{2,\infty} \} \lesssim \frac{\sigma\kappa\sqrt{n_2\mu r/n_1} + \sigma\sqrt{r\log n}}{\sigma_r^*}.$$

We can then repeat the same analysis as in the proof of Lemma 4.17 to connect $\|\mathbf{U}\mathbf{H}_U - \mathbf{U}^*\|_{2,\infty}$ (resp. $\|\mathbf{V}\mathbf{H}_V - \mathbf{V}^*\|_{2,\infty}$) with $\|\mathbf{U}\text{sgn}(\mathbf{H}_U) - \mathbf{U}^*\|_{2,\infty}$ (resp. $\|\mathbf{V}\text{sgn}(\mathbf{H}_V) - \mathbf{V}^*\|_{2,\infty}$), and obtain the desired bound in (4.41); the details are omitted here for the sake of conciseness.

We now proceed to the second claim (4.42). Towards this, invoke Corollary 4.3 and the inequality (4.145) to obtain

$$\|\bar{\mathbf{U}}\bar{\Lambda}\bar{\mathbf{U}}^\top - \mathcal{S}(\mathbf{M}^*)\|_\infty \lesssim \sigma\kappa^2\bar{\mu}r\sqrt{\frac{\log n}{n}} \lesssim \sigma\kappa^2\mu r\sqrt{\frac{n\log n}{n_1^2}}.$$

This in conjunction with the following observations

$$\bar{\mathbf{U}}\bar{\Lambda}\bar{\mathbf{U}}^\top = \mathcal{S}(\mathbf{U}\Sigma\mathbf{V}^\top)$$

$$\|\mathcal{S}(\mathbf{U}\Sigma\mathbf{V}^\top) - \mathcal{S}(\mathbf{M}^*)\|_\infty = \|\mathbf{U}\Sigma\mathbf{V}^\top - \mathbf{M}^*\|_\infty$$

immediately establishes the second claim.

4.11 Appendix D: Proof of Theorem 4.10

As before (see (4.85)), we assume throughout the proof that $\lambda_r^* > 0$ for the purpose of simplifying notation.

4.11.1 Proof outline

We now outline the proof of our distributional guarantees in Theorem 4.10. The first step consists of justifying the heuristic first-order approximation in (4.65). This is stated in the lemma below, with the proof deferred to Section 4.11.2.

Lemma 4.18 (First-order approximation). Suppose that the assumptions of Theorem 4.2 hold. Then with probability at least $1 - O(n^{-5})$, one can write

$$\mathbf{U}\text{sgn}(\mathbf{H}) - \mathbf{U}^* = \underbrace{\mathbf{E}\mathbf{U}^*(\Lambda^*)^{-1}}_{=: \mathbf{Z}} + \mathbf{\Psi}, \quad (4.147a)$$

$$\mathbf{M} - \mathbf{M}^* = \underbrace{\mathbf{E}\mathbf{U}^*\mathbf{U}^{*\top} + \mathbf{U}^*\mathbf{U}^{*\top}\mathbf{E}}_{=: \mathbf{W}} + \mathbf{\Phi} \quad (4.147b)$$

for some matrices $\mathbf{\Psi}$ and $\mathbf{\Phi}$ obeying

$$\|\mathbf{\Psi}\|_{2,\infty} \lesssim \frac{\sigma^2 \kappa \sqrt{\mu r n \log^2 n}}{(\lambda_r^*)^2} + \frac{\kappa \sigma^2 \sqrt{\mu r n}}{(\lambda_r^*)^2} + \frac{\sigma}{\lambda_r^*} \sqrt{\frac{\mu r^2 \log n}{n}}, \quad (4.148a)$$

$$\|\mathbf{\Phi}\|_\infty \lesssim \frac{\sigma^2 \mu r \kappa^2 \log n}{\lambda_r^*} + \frac{\sigma \kappa \mu \sqrt{r^3 \log n}}{n}. \quad (4.148b)$$

Remark 4.8. In addition to quantifying the goodness of the approximation (4.147b), Lemma 4.18 also delivers a more refined characterization for the first-order approximation $\mathbf{U}\text{sgn}(\mathbf{H}) - \mathbf{U}^* \approx \mathbf{E}\mathbf{U}^*(\Lambda^*)^{-1}$ in comparison to Theorem 4.2. As it turns out, this result (4.147a) also assists in performing statistical inference on the low-rank factors \mathbf{U}^* . The interested reader is referred to Yan *et al.* (2021) for details.

In turn, Lemma 4.18 motivates one to pin down the distribution of the matrix \mathbf{W} in (4.147b). This can be accomplished by invoking the Berry-Esseen Theorem (e.g., Chen *et al.* (2010, Theorem 3.7)), which gives rise to the following distributional characterization. The proof of this lemma can be found in Section 4.11.3.

Lemma 4.19 (Gaussian approximation). Suppose that the assumptions of Theorem 4.2 hold, and that

$$\frac{\|\mathbf{U}_{j,\cdot}^*\|_2^2 + \|\mathbf{U}_{i,\cdot}^*\|_2^2}{\|\mathbf{U}^*\|_{\text{F}}^2} \gtrsim \frac{B^2 \kappa^2 \mu^2 r^2 \log^2 n}{\sigma_{\min}^2 n^2} + \frac{\sigma^4 \mu^2 r \kappa^4 \log^3 n}{\sigma_{\min}^2 (\lambda_r^*)^2}. \quad (4.149)$$

Let $\mathbf{W} = \mathbf{E} \mathbf{U}^* \mathbf{U}^{*\top} + \mathbf{U}^* \mathbf{U}^{*\top} \mathbf{E}$. For any $1 \leq i, j \leq n$, one has

$$\sup_{z \in \mathbb{R}} \left| \mathbb{P} \left(W_{i,j} \leq z \sqrt{v_{i,j}^*} \right) - \Phi(z) \right| = o(1), \quad (4.150a)$$

$$\|\Phi\|_{\infty} = o \left(\sqrt{v_{i,j}^*} \right), \quad (4.150b)$$

where $v_{i,j}^*$ is defined in (4.68), and $\Phi(\cdot)$ denotes the CDF of the standard Gaussian distribution.

To finish up, invoke Lemma 4.18 and (4.150b) to yield

$$M_{i,j} - M_{i,j}^* = W_{i,j} + \delta_{\phi} \sqrt{v_{i,j}^*} \quad \text{with } \delta_{\phi} = o(1).$$

A little algebra further gives

$$\begin{aligned} \left| \mathbb{P} \left(M_{i,j} - M_{i,j}^* \leq z \sqrt{v_{i,j}^*} \right) - \Phi(z) \right| &= \left| \mathbb{P} \left(W_{i,j} \leq (z - \delta_{\phi}) \sqrt{v_{i,j}^*} \right) - \Phi(z) \right| \\ &\leq \left| \mathbb{P} \left(W_{i,j} \leq (z - \delta_{\phi}) \sqrt{v_{i,j}^*} \right) - \Phi(z - \delta_{\phi}) \right| + |\Phi(z - \delta_{\phi}) - \Phi(z)| \\ &\leq o(1) + \delta_{\phi} = o(1), \end{aligned}$$

where the last line invokes Lemma 4.19 and the fact that $|\Phi(u) - \Phi(v)| \leq |u - v|$ for any $u, v \in \mathbb{R}$. This completes the proof of Theorem 4.10, as long as Lemmas 4.18 and 4.19 can be established. The rest of this section is thus devoted to proving Lemmas 4.18 and 4.19.

4.11.2 Proof of Lemma 4.18

Before proceeding to the proof, we make note of several preliminary facts that are all direct consequences of the analysis of Theorem 4.2

and Corollary 4.3. The proof of these preliminary results can be found in Section 4.11.4.

Lemma 4.20. With probability exceeding $1 - O(n^{-5})$, one can write

$$\mathbf{U}\Lambda \text{sgn}(\mathbf{H}) = \mathbf{M}\mathbf{U}^* + \Delta_1, \quad (4.151a)$$

$$\text{sgn}(\mathbf{H})\Lambda^* \text{sgn}(\mathbf{H})^\top = \Lambda + \Delta_2 \quad (4.151b)$$

for some matrices Δ_1 and Δ_2 obeying

$$\|\Delta_1\|_{2,\infty} \lesssim \frac{\sigma^2 \kappa \sqrt{\mu r n \log^2 n}}{\lambda_r^*}, \quad (4.152a)$$

$$\|\Delta_2\| \lesssim \frac{\kappa \sigma^2 n}{\lambda_r^*} + \sigma \sqrt{r \log n}. \quad (4.152b)$$

We are now ready to embark on the proof of Lemma 4.18. In order to analyze the behavior of $\mathbf{U}\text{sgn}(\mathbf{H})$, we first point out the following decomposition:

$$\begin{aligned} \mathbf{U}\text{sgn}(\mathbf{H})\Lambda^* &= \mathbf{U}\Lambda \text{sgn}(\mathbf{H}) + \mathbf{U}(\text{sgn}(\mathbf{H})\Lambda^* - \Lambda \text{sgn}(\mathbf{H})) \\ &= \mathbf{M}\mathbf{U}^* + \Delta_1 + \mathbf{U}(\text{sgn}(\mathbf{H})\Lambda^* \text{sgn}(\mathbf{H})^\top - \Lambda) \text{sgn}(\mathbf{H}) \\ &= \mathbf{M}^* \mathbf{U}^* + \mathbf{E}\mathbf{U}^* + \Delta_1 + \mathbf{U}\Delta_2 \text{sgn}(\mathbf{H}) \\ &= \mathbf{U}^* \Lambda^* + \mathbf{E}\mathbf{U}^* + \Delta_1 + \mathbf{U}\Delta_2 \text{sgn}(\mathbf{H}), \end{aligned}$$

where the second and the third identities result from Lemma 4.20 (cf. (4.151)). The key point of this decomposition is to establish a connection between $\mathbf{U}\text{sgn}(\mathbf{H})\Lambda^*$ and $\mathbf{U}^* \Lambda^* + \mathbf{E}\mathbf{U}^*$, with the assistance of the matrices Δ_1 and Δ_2 studied in Lemma 4.20. A little algebra then yields

$$\mathbf{U}\text{sgn}(\mathbf{H}) - \mathbf{U}^* = \underbrace{\mathbf{E}\mathbf{U}^* (\Lambda^*)^{-1}}_{=: \mathbf{Z}} + \underbrace{\Delta_1 (\Lambda^*)^{-1} + \mathbf{U}\Delta_2 \text{sgn}(\mathbf{H}) (\Lambda^*)^{-1}}_{=: \mathbf{\Psi}}. \quad (4.153)$$

In view of Lemma 4.20, the residual matrix $\mathbf{\Psi}$ obeys

$$\begin{aligned} \|\mathbf{\Psi}\|_{2,\infty} &\leq \|\Delta_1\|_{2,\infty} \|(\Lambda^*)^{-1}\| + \|\mathbf{U}\|_{2,\infty} \|\Delta_2\| \|\text{sgn}(\mathbf{H})\| \|(\Lambda^*)^{-1}\| \\ &\lesssim \frac{1}{\lambda_r^*} \|\Delta_1\|_{2,\infty} + \frac{1}{\lambda_r^*} \sqrt{\frac{\mu r}{n}} \|\Delta_2\| \end{aligned}$$

$$\lesssim \frac{\sigma^2 \kappa \sqrt{\mu r n \log^2 n}}{(\lambda_r^*)^2} + \frac{\kappa \sigma^2 \sqrt{\mu r n}}{(\lambda_r^*)^2} + \frac{\sigma}{\lambda_r^*} \sqrt{\frac{\mu r^2 \log n}{n}} \quad (4.154)$$

as claimed.

The next step lies in analyzing the matrix estimator $\mathbf{M} = \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^\top$. Towards this, we make the observation that

$$\begin{aligned} \mathbf{M} - \mathbf{M}^* &= \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^\top - \mathbf{M}^* \\ &= \mathbf{U} (\text{sgn}(\mathbf{H}) \boldsymbol{\Lambda}^* \text{sgn}(\mathbf{H})^\top) \mathbf{U}^\top - \mathbf{U} \boldsymbol{\Delta}_2 \mathbf{U}^\top - \mathbf{M}^* \\ &= (\mathbf{U}^* + \mathbf{Z} + \boldsymbol{\Psi}) \boldsymbol{\Lambda}^* (\mathbf{U}^* + \mathbf{Z} + \boldsymbol{\Psi})^\top - \mathbf{U} \boldsymbol{\Delta}_2 \mathbf{U}^\top - \mathbf{U}^* \boldsymbol{\Lambda}^* \mathbf{U}^{*\top} \\ &= \mathbf{Z} \boldsymbol{\Lambda}^* \mathbf{U}^{*\top} + \mathbf{U}^* \boldsymbol{\Lambda}^* \mathbf{Z}^\top + \boldsymbol{\Phi} \\ &= \mathbf{E} \mathbf{U}^* \mathbf{U}^{*\top} + \mathbf{U}^* \mathbf{U}^{*\top} \mathbf{E} + \boldsymbol{\Phi}, \end{aligned} \quad (4.155)$$

where the second line relies on Lemma 4.20 (cf. (4.151b)), the third identity makes use of (4.153), and the residual matrix $\boldsymbol{\Phi}$ is defined as

$$\boldsymbol{\Phi} := \boldsymbol{\Psi} \boldsymbol{\Lambda}^* (\mathbf{U} \text{sgn}(\mathbf{H}))^\top + \mathbf{U}^* \boldsymbol{\Lambda}^* \boldsymbol{\Psi}^\top + \mathbf{Z} \boldsymbol{\Lambda}^* (\mathbf{U} \text{sgn}(\mathbf{H}) - \mathbf{U}^*)^\top - \mathbf{U} \boldsymbol{\Delta}_2 \mathbf{U}^\top. \quad (4.156)$$

In addition, it is seen from (4.92) that

$$\|\mathbf{Z}\|_{2,\infty} \leq \|\mathbf{E} \mathbf{U}^* (\boldsymbol{\Lambda}^*)^{-1}\|_{2,\infty} \leq \frac{1}{\lambda_r^*} \|\mathbf{E} \mathbf{U}^*\|_{2,\infty} \lesssim \frac{\sigma \sqrt{r \log n}}{\lambda_r^*}. \quad (4.157)$$

Consequently, one can deduce that

$$\begin{aligned} \|\boldsymbol{\Phi}\|_\infty &\leq \|\boldsymbol{\Lambda}^*\| \left\{ \|\mathbf{Z}\|_{2,\infty}^2 + \|\mathbf{Z}\|_{2,\infty} \|\mathbf{U} \text{sgn}(\mathbf{H}) - \mathbf{U}^*\|_{2,\infty} \right\} \\ &\quad + \|\boldsymbol{\Lambda}^*\| \|\boldsymbol{\Psi}\|_{2,\infty} \left(\|\mathbf{U}\|_{2,\infty} + \|\mathbf{U}^*\|_{2,\infty} \right) + \|\mathbf{U}\|_{2,\infty}^2 \|\boldsymbol{\Delta}_2\| \\ &\lesssim |\lambda_1^*| \left\{ \frac{\sigma^2 r \log n}{(\lambda_r^*)^2} + \frac{\sigma \sqrt{r \log n}}{\lambda_r^*} \cdot \frac{\sigma \kappa \sqrt{\mu r \log n}}{\lambda_r^*} \right\} \\ &\quad + |\lambda_1^*| \sqrt{\frac{\mu r}{n}} \left\{ \frac{\sigma^2 \kappa \sqrt{\mu r n \log^2 n}}{(\lambda_r^*)^2} + \frac{\kappa \sigma^2 \sqrt{\mu r n}}{(\lambda_r^*)^2} + \frac{\sigma}{\lambda_r^*} \sqrt{\frac{\mu r^2 \log n}{n}} \right\} \\ &\quad + \frac{\mu r}{n} \left\{ \frac{\kappa \sigma^2 n}{\lambda_r^*} + \sigma \sqrt{r \log n} \right\} \\ &\lesssim \frac{\sigma^2 \mu r \kappa^2 \log n}{\lambda_r^*} + \frac{\sigma \kappa \mu \sqrt{r^3 \log n}}{n}, \end{aligned}$$

where the second inequality follows from (4.154), (4.157), (4.133), Theorem 4.2, and Lemma 4.20.

4.11.3 Proof of Lemma 4.19

In what follows, we shall only focus on the case with $i \neq j$. The case with $i = j$ can be analyzed in an analogous manner; we omit it for the sake of brevity. Before proceeding to the proof, we make note of a couple of basic facts about \mathbf{P}^* that will prove useful. The first property asserts that, for any $1 \leq j \leq n$,

$$\begin{aligned} \sum_{l=1}^n P_{j,l}^{*2} &= \sum_{l=1}^n P_{l,j}^{*2} = \|(\mathbf{U}^* \mathbf{U}^{*\top})_{\cdot,j}\|_2^2 = \|\mathbf{U}^* (\mathbf{U}_{j,\cdot}^*)^\top\|_2^2 = \mathbf{U}_{j,\cdot}^* \mathbf{U}^{*\top} \mathbf{U}^* (\mathbf{U}_{j,\cdot}^*)^\top \\ &= \|\mathbf{U}_{j,\cdot}^*\|_2^2 = P_{j,j}^*. \end{aligned} \quad (4.158)$$

The second property is concerned with the term $\|\mathbf{P}^*\|_\infty$:

$$\|\mathbf{P}^*\|_\infty = \|\mathbf{U}^* \mathbf{U}^{*\top}\|_\infty \leq \|\mathbf{U}^*\|_{2,\infty}^2 \leq \frac{\mu r}{n}, \quad (4.159)$$

where the last inequality follows from the incoherence assumption.

In view of the definition (4.66) of \mathbf{P}^* , we can express $\mathbf{W} = \mathbf{E} \mathbf{P}^* + \mathbf{P}^* \mathbf{E}$, which reveals that

$$W_{i,j} = \sum_{l: l \neq j} E_{i,l} P_{l,j}^* + \sum_{l: l \neq i} P_{i,l}^* E_{l,j} + E_{i,j} (P_{i,i}^* + P_{j,j}^*). \quad (4.160)$$

In other words, $W_{i,j}$ can be viewed as a weighted sum of independent random variables $\{E_{i,l} \mid l \neq j\} \cup \{E_{l,j} \mid l \neq i\} \cup \{E_{i,j}\}$. To pin down the distribution of $W_{i,j}$, we resort to a non-asymptotic version of the celebrated Berry-Esseen Theorem; see Chen *et al.* (2010, Theorem 3.7) for a proof using Stein's method.

Theorem 4.21 (The Berry-Esseen bound). Let ξ_1, \dots, ξ_n be independent zero-mean random variables satisfying $\sum_{i=1}^n \text{Var}(\xi_i) = v$. Then the quantity $S = \frac{1}{\sqrt{v}} \sum_{i=1}^n \xi_i$ satisfies

$$\sup_{z \in \mathbb{R}} |\mathbb{P}(S \leq z) - \Phi(z)| \leq 10\gamma, \quad \text{where } \gamma = \sum_{i=1}^n \frac{\mathbb{E}[|\xi_i|^3]}{v^{3/2}}.$$

According to the Berry-Esseen bound (cf. Theorem 4.21), proving the approximate Gaussianity of $W_{i,j}$ boils down to characterizing the second and the third moments of these random variables under consideration.

Let us start with the variance statistics. Given that $\{E_{i,j} \mid i \geq j\}$ are independently generated, we can straightforwardly see that

$$\begin{aligned} v_{i,j}^* &= \sum_{l: l \neq j} \sigma_{i,l}^2 P_{l,j}^{*2} + \sum_{l: l \neq i} P_{i,l}^{*2} \sigma_{l,j}^2 + \sigma_{i,j}^2 (P_{i,i}^* + P_{j,j}^*)^2 \\ &= \text{Var}(W_{i,j}). \end{aligned} \quad (4.161)$$

We now develop a lower bound on this variance term. Given that $\mathbf{P}^* \succeq \mathbf{0}$, one has $P_{i,i}^*, P_{j,j}^* \geq 0$, which combined with (4.158) reveals that

$$v_{i,j}^* \geq \sigma_{\min}^2 \left\{ \sum_{l=1}^n P_{l,j}^{*2} + \sum_{l=1}^n P_{i,l}^{*2} \right\} = \sigma_{\min}^2 \left\{ \|\mathbf{U}_{j,\cdot}^*\|_2^2 + \|\mathbf{U}_{i,\cdot}^*\|_2^2 \right\}. \quad (4.162)$$

Next, we move on to bound the third moments. Utilizing the independence of $\{E_{i,j} \mid i \geq j\}$ once again gives

$$\begin{aligned} \gamma &:= \frac{\sum_{l: l \neq j} \mathbb{E}[|E_{i,l}|^3] |P_{l,j}^*|^3 + \sum_{l: l \neq i} |P_{i,l}|^3 \mathbb{E}[|E_{l,j}|^3] + \mathbb{E}[|E_{i,j}|^3] |P_{i,i}^* + P_{j,j}^*|^3}{(v_{i,j}^*)^{3/2}} \\ &\leq \frac{2B\|\mathbf{P}^*\|_\infty}{(v_{i,j}^*)^{3/2}} \left\{ \sum_{l: l \neq j} \mathbb{E}[E_{i,l}^2] P_{l,j}^{*2} + \sum_{l: l \neq i} P_{i,l}^{*2} \mathbb{E}[E_{l,j}^2] + \mathbb{E}[E_{i,j}^2] |P_{i,i}^* + P_{j,j}^*|^2 \right\} \\ &= \frac{2B\|\mathbf{P}^*\|_\infty}{(v_{i,j}^*)^{3/2}} \cdot v_{i,j}^* = \frac{2B\|\mathbf{P}^*\|_\infty}{(v_{i,j}^*)^{1/2}}, \end{aligned} \quad (4.163)$$

where we have used the assumption that $|E_{i,j}| \leq B$, and the last line arises from the expression (4.161). Substituting (4.159) and (4.162) into (4.163) and invoking the elementary identity $\|\mathbf{U}^*\|_{\text{F}}^2 = r$, we arrive at

$$\gamma \leq \frac{2B\mu r}{n\sigma_{\min} \sqrt{\|\mathbf{U}_{j,\cdot}^*\|_2^2 + \|\mathbf{U}_{i,\cdot}^*\|_2^2}} = \frac{2B\mu\sqrt{r} \|\mathbf{U}^*\|_{\text{F}}}{n\sigma_{\min} \sqrt{\|\mathbf{U}_{j,\cdot}^*\|_2^2 + \|\mathbf{U}_{i,\cdot}^*\|_2^2}} = o(1),$$

provided that Condition (4.149) holds.

With the above calculations in place, invoking Theorem 4.21 immediately leads to

$$\sup_{z \in \mathbb{R}} \left| \mathbb{P}(W_{i,j} \leq \sqrt{v_{i,j}^*} z) - \Phi(z) \right| \leq 10\gamma = o(1).$$

as claimed in (4.150a).

Finally, we turn to proving the bound (4.150b). By virtue of Lemma 4.18 and (4.162), we know that

$$\begin{aligned}\|\Phi\|_\infty &\lesssim \frac{\sigma^2 \mu r \kappa^2 \log n}{\lambda_r^*} + \frac{\sigma \kappa \mu \sqrt{r^3 \log n}}{n} \\ &= o\left(\sigma_{\min} \sqrt{\|\mathbf{U}_{j,\cdot}^*\|_2^2 + \|\mathbf{U}_{i,\cdot}^*\|_2^2}\right) \leq o(\sqrt{v_{i,j}^*}),\end{aligned}$$

with the proviso that

$$\begin{aligned}\sqrt{\|\mathbf{U}_{j,\cdot}^*\|_2^2 + \|\mathbf{U}_{i,\cdot}^*\|_2^2} &\gtrsim \frac{\sigma^2 \mu r \kappa^2 \log^{3/2} n}{\sigma_{\min} \lambda_r^*} + \frac{\sigma \kappa \mu \sqrt{r^3} \log n}{\sigma_{\min} n} \\ &= \left(\frac{\sigma^2 \mu \sqrt{r} \kappa^2 \log^{3/2} n}{\sigma_{\min} \lambda_r^*} + \frac{\sigma \kappa \mu r \log n}{\sigma_{\min} n}\right) \|\mathbf{U}^*\|_{\text{F}}. \quad (4.164)\end{aligned}$$

Recognizing the trivial bound $\sigma^2 = \max_{i,j} \mathbb{E}[E_{i,j}^2] \leq B^2$, we know that Condition (4.164) holds as long as

$$\frac{\|\mathbf{U}_{j,\cdot}^*\|_2^2 + \|\mathbf{U}_{i,\cdot}^*\|_2^2}{\|\mathbf{U}^*\|_{\text{F}}^2} \gtrsim \frac{\sigma^4 \mu^2 r \kappa^4 \log^3 n}{\sigma_{\min}^2 (\lambda_r^*)^2} + \frac{B^2 \kappa^2 \mu^2 r^2 \log^2 n}{\sigma_{\min}^2 n^2}, \quad (4.165)$$

which is precisely Condition (4.149). This concludes the proof of Lemma 4.19.

4.11.4 Proof of Lemma 4.20

To begin with, let us begin by proving (4.152a). From the definition of the quantity \mathcal{E}_1 (see Lemma 4.16), we have

$$\begin{aligned}\|\mathbf{M}(\mathbf{UH} - \mathbf{U}^*)\|_{2,\infty} &= \lambda_r^* \mathcal{E}_1 / 2 \\ &\lesssim (\alpha_0 + \alpha_1 + \alpha_2) + (\sigma \sqrt{n} + B \log n) \|\mathbf{UH} - \mathbf{U}^*\|_{2,\infty} \\ &\lesssim (\alpha_0 + \alpha_1 + \alpha_2) + (\sigma \sqrt{n} + B \log n) \frac{\sigma \kappa \sqrt{\mu r \log n}}{\lambda_r^*} \\ &\asymp \frac{\sigma^2 \kappa \sqrt{\mu r n \log^2 n}}{\lambda_r^*},\end{aligned}$$

where the first inequality comes from (4.116) and (4.117), the second inequality is a consequence of Theorem 4.2, and the last line relies on

our previous bounds on $\alpha_0, \alpha_1, \alpha_2$ (see (4.112), (4.113) and (4.115)) and holds as long as $B \lesssim \sigma\sqrt{n}/(\mu \log n)$. Additionally, from the elementary identity $\mathbf{M}\mathbf{U}\mathbf{H} = \mathbf{U}\Lambda\mathbf{H}$, we obtain

$$\begin{aligned}\|\mathbf{U}\Lambda\text{sgn}(\mathbf{H}) - \mathbf{M}\mathbf{U}\mathbf{H}\|_{2,\infty} &= \|\mathbf{U}\Lambda\text{sgn}(\mathbf{H}) - \mathbf{U}\Lambda\mathbf{H}\|_{2,\infty} \\ &\leq \|\mathbf{U}\|_{2,\infty} \|\Lambda\| \|\text{sgn}(\mathbf{H}) - \mathbf{H}\| \\ &\lesssim \sqrt{\frac{\mu r}{n}} |\lambda_1^*| \cdot \frac{\sigma^2 n}{(\lambda_r^*)^2} = \frac{\kappa \sigma^2 \sqrt{\mu r n}}{\lambda_r^*},\end{aligned}$$

where the last line results from Lemma 4.15, the fact (4.133), and the following inequality

$$\|\Lambda\| \leq \|\Lambda^*\| + \|\mathbf{E}\| \leq |\lambda_1^*| + O(\sigma\sqrt{n}) \leq 2|\lambda_1^*|. \quad (4.166)$$

Taking together the above bounds and applying the triangle inequality immediately establish (4.152a).

Next, we turn to the proof of the bound (4.152b). Note that it has been shown in (4.139) that

$$\|\mathbf{H}\Lambda^*\mathbf{H}^\top - \Lambda\| \lesssim \frac{\kappa \sigma^2 n}{\lambda_r^*} + \sigma\sqrt{r \log n}.$$

In addition, the triangle inequality leads to

$$\begin{aligned}\|\text{sgn}(\mathbf{H})\Lambda^*\text{sgn}(\mathbf{H})^\top - \mathbf{H}\Lambda^*\mathbf{H}^\top\| &\leq \|\text{sgn}(\mathbf{H})\Lambda^*(\text{sgn}(\mathbf{H}) - \mathbf{H})^\top\| + \|(\text{sgn}(\mathbf{H}) - \mathbf{H})\Lambda^*\mathbf{H}^\top\| \\ &\leq (\|\text{sgn}(\mathbf{H})\| + \|\mathbf{H}\|) \|\Lambda^*\| \|\text{sgn}(\mathbf{H}) - \mathbf{H}\| \\ &\leq 2|\lambda_1^*| \|\text{sgn}(\mathbf{H}) - \mathbf{H}\| \\ &\lesssim |\lambda_1^*| \frac{\sigma^2 n}{(\lambda_r^*)^2} = \frac{\kappa \sigma^2 n}{\lambda_r^*},\end{aligned}$$

where the third line follows since $\|\text{sgn}(\mathbf{H})\| = 1$ and $\|\mathbf{H}\| \leq \|\mathbf{U}\| \|\mathbf{U}^*\| = 1$, and the last inequality comes from (4.123). Combining the above two results and invoking the triangle inequality lead to the advertised bound (4.152b).

4.12 Appendix E: Proof of Theorem 4.11

With the distributional guarantees in Theorem 4.10 in place, the only remaining task boils down to verifying the statistical accuracy of the

variance estimator $\hat{v}_{i,j}$. This can be achieved via the following lemma, whose proof is provided in Section 4.12.1.

Lemma 4.22. Suppose that the assumptions of Theorem 4.2 hold. In addition, assume that $\kappa^4 \mu^2 r^2 \log n \leq n$, $\sigma \sqrt{n} \lesssim |\lambda_r^*|/\kappa$ and

$$\|\mathbf{U}_{j,\cdot}^*\|_2^2 + \|\mathbf{U}_{i,\cdot}^*\|_2^2 \gtrsim \frac{B\sigma\kappa^2\mu^2r^2\log n}{\sigma_{\min}^2 n^{3/2}} + \frac{\sigma^3\kappa\mu^2r^2\log n}{\sigma_{\min}^2 |\lambda_r^*| \sqrt{n}}. \quad (4.167)$$

With probability exceeding $1 - O(n^{-5})$, one has

$$|\hat{v}_{i,j} - v_{i,j}^*| = o(v_{i,j}^*). \quad (4.168)$$

Lemma 4.22 essentially enables us to express

$$\hat{v}_{i,j} = (1 + \zeta_v)^2 v_{i,j}^* \quad \text{with } \zeta_v = o(1).$$

As a consequence, we can further demonstrate that

$$\begin{aligned} |\mathbb{P}(M_{i,j}^* \in \text{CI}_{i,j}^{1-\alpha}) - (1 - \alpha)| &= \left| \mathbb{P}\left(\left|\hat{M}_{i,j} - M_{i,j}^*\right| \leq z_{\alpha/2} \sqrt{\hat{v}_{i,j}}\right) - (1 - \alpha) \right| \\ &= \left| \mathbb{P}\left(\left|\hat{M}_{i,j} - M_{i,j}^*\right| \leq (1 + \zeta_v) z_{\alpha/2} \sqrt{v_{i,j}^*}\right) - (1 - \alpha) \right| \\ &\leq |\Phi((1 + \zeta_v) z_{\alpha/2}) - \Phi(-(1 + \zeta_v) z_{\alpha/2}) - (1 - \alpha)| + o(1) \\ &\leq |\Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2}) - (1 - \alpha)| + 2 |\Phi((1 + \zeta_v) z_{\alpha/2}) - \Phi(z_{\alpha/2})| + o(1) \\ &\leq 2\zeta_v z_{\alpha/2} = o(1), \end{aligned}$$

where the first inequality follows from Theorem 4.10 and $z_{\alpha/2} := \Phi^{-1}(1 - \alpha/2)$, the second inequality applies the triangle inequality, and the validity of the last line can be seen from the basic fact $|\Phi(u) - \Phi(v)| \leq |u - v|$.

When $\sigma_{\min} \asymp \sigma$, Condition (4.167) simplifies to

$$\frac{\|\mathbf{U}_{j,\cdot}^*\|_2^2 + \|\mathbf{U}_{i,\cdot}^*\|_2^2}{\|\mathbf{U}^*\|_{\text{F}}^2} \gtrsim \frac{B\kappa^2\mu^2r^2\log n}{\sigma n^{3/2}} + \frac{\sigma\kappa\mu^2r^2\log n}{|\lambda_r^*|\sqrt{n}}. \quad (4.169)$$

We still need to ensure that Condition (4.69) is satisfied. It is seen that

$$\frac{B^2\kappa^2\mu^2r^2\log^2 n}{\sigma^2 n^2} + \frac{\sigma^2\mu^2r\kappa^4\log^3 n}{(\lambda_r^*)^2} \lesssim \frac{B\kappa^2\mu^2r^2\log^2 n}{\sigma n^{3/2}} + \frac{\sigma\mu^2r\kappa^3\log^3 n}{|\lambda_r^*|\sqrt{n}},$$

which holds if $\sigma\kappa\sqrt{n} \lesssim |\lambda_r^*|$ and $B \lesssim \sigma\sqrt{n}$. As a result, if Condition (4.74) holds, then both (4.169) and (4.69) are satisfied. This finishes the proof, as long as Lemma 4.22 can be established.

4.12.1 Proof for Lemma 4.22

As before, we shall only present the proof for the case with $i \neq j$ for the sake of conciseness. In order to justify the goodness of the estimator $\widehat{v}_{i,j}$, we find it convenient to first look at the surrogate estimator introduced in (4.70), i.e.,

$$\tilde{v}_{i,j} = \sum_{l=1}^n E_{i,l}^2 P_{l,j}^{*2} + \sum_{l=1}^n P_{i,l}^{*2} E_{l,j}^2 + 2E_{i,j}^2 P_{i,i}^* P_{j,j}^*, \quad (4.170)$$

In the sequel, our proof consists of two main steps:

- Show that the surrogate $\tilde{v}_{i,j}$ is a reliable estimate of the truth $v_{i,j}^*$, namely, $\tilde{v}_{i,j} \approx v_{i,j}^*$.
- Show that the estimator in use and the surrogate estimator are sufficiently close, namely, $\widehat{v}_{i,j} \approx \tilde{v}_{i,j}$.

Step 1: show that $\tilde{v}_{i,j} \approx v_{i,j}^*$

Firstly, the fact that the $E_{i,j}$'s are zero-mean random variables immediately reveals that $\tilde{v}_{i,j}$ is an unbiased estimate of $v_{i,j}^*$, that is,

$$\mathbb{E}[\tilde{v}_{i,j}] = v_{i,j}^*.$$

Secondly, given that the $E_{i,j}$'s are statistically independent, we intend to invoke the Bernstein inequality to control the difference $\tilde{v}_{i,j} - v_{i,j}^* = \tilde{v}_{i,j} - \mathbb{E}[\tilde{v}_{i,j}]$. To do so, one first calculates that

$$L_0 := \max \left\{ \max_l E_{i,l}^2 P_{l,j}^{*2}, \max_l E_{l,j}^2 P_{i,l}^{*2}, E_{i,j}^2 (P_{i,i}^* + P_{j,j}^*)^2 \right\} \leq 4B^2 \|\mathbf{P}^*\|_\infty^2$$

and

$$\begin{aligned} V_0 &:= \sum_{l: l \neq j} \text{Var}(E_{i,l}^2) P_{l,j}^{*4} + \sum_{l: l \neq i} \text{Var}(E_{l,j}^2) P_{i,l}^{*4} + \text{Var}(E_{i,j}^2) (P_{i,i}^* + P_{j,j}^*)^4 \\ &\lesssim \sum_{l=1}^n \mathbb{E}[E_{i,l}^4] P_{l,j}^{*4} + \sum_{l=1}^n \mathbb{E}[E_{l,j}^4] P_{i,l}^{*4} + \mathbb{E}[E_{i,j}^4] P_{i,i}^{*2} P_{j,j}^{*2} \\ &\lesssim B^2 \|\mathbf{P}^*\|_\infty^2 \left\{ \sum_{l=1}^n \mathbb{E}[E_{i,l}^2] P_{l,j}^{*2} + \sum_{l=1}^n \mathbb{E}[E_{l,j}^2] P_{i,l}^{*2} + \mathbb{E}[E_{i,j}^2] P_{i,i}^* P_{j,j}^* \right\} \end{aligned}$$

$$\begin{aligned} &\lesssim \sigma^2 B^2 \|\mathbf{P}^*\|_\infty^2 \left\{ \sum_{l=1}^n P_{l,j}^{*2} + \sum_{l=1}^n P_{i,l}^{*2} \right\} \\ &\lesssim \sigma^2 B^2 \|\mathbf{P}^*\|_\infty^3, \end{aligned}$$

where the last line follows from (4.158). Invoking the Bernstein inequality (cf. Corollary 3.3) reveals that with probability exceeding $1 - O(n^{-5})$,

$$\begin{aligned} |\tilde{v}_{i,j} - v_{i,j}^*| &= |\tilde{v}_{i,j} - \mathbb{E}[\tilde{v}_{i,j}]| \lesssim \sqrt{V_0 \log n} + L_0 \log n \\ &\lesssim \sigma B \sqrt{\|\mathbf{P}^*\|_\infty^3 \log n} + B^2 \|\mathbf{P}^*\|_\infty^2 \log n \\ &\lesssim \frac{\sigma B \mu^{3/2} r^{3/2} \sqrt{\log n}}{n^{3/2}} + \frac{\mu^2 r^2 B^2 \log n}{n^2}, \end{aligned} \quad (4.171)$$

where the last inequality results from (4.159).

Step 2: show that $\hat{v}_{i,j} \approx \tilde{v}_{i,j}$

In order to accomplish this, we are in need of controlling the difference between $E_{i,j}$ (resp. $P_{i,j}^*$) and $\hat{E}_{i,j}$ (resp. $\hat{P}_{i,j}$). To this end, apply the entrywise estimation guarantees in Corollary 4.3 to yield

$$\begin{aligned} \|\hat{\mathbf{E}} - \mathbf{E}\|_\infty &= \|(\mathbf{M} - \mathbf{U}\Lambda\mathbf{U}^\top) - (\mathbf{M} - \mathbf{M}^*)\|_\infty \\ &= \|\mathbf{U}\Lambda\mathbf{U}^\top - \mathbf{M}^*\|_\infty \lesssim \sigma \kappa^2 \mu r \sqrt{\frac{\log n}{n}}, \end{aligned} \quad (4.172)$$

and as a result,

$$\|\hat{\mathbf{E}}\|_\infty \leq \|\mathbf{E}\|_\infty + \|\hat{\mathbf{E}} - \mathbf{E}\|_\infty \lesssim B + \sigma \kappa^2 \mu r \sqrt{\frac{\log n}{n}} \asymp B, \quad (4.173)$$

provided that $B \gtrsim \sigma \kappa^2 \mu r \sqrt{\frac{\log n}{n}}$ (which is trivially satisfied if $\kappa^2 \mu r \sqrt{\frac{\log n}{n}} \leq 1$). Moving on to the error term $\hat{P}_{i,j} - P_{i,j}^*$, we observe that

$$\begin{aligned} \|\hat{\mathbf{P}} - \mathbf{P}^*\|_\infty &= \|\mathbf{U}\mathbf{U}^\top - \mathbf{U}^*\mathbf{U}^{*\top}\|_\infty \\ &\leq \|(\mathbf{U}\text{sgn}(\mathbf{H}) - \mathbf{U}^*)\mathbf{U}^{*\top}\|_\infty + \|\mathbf{U}\text{sgn}(\mathbf{H})(\mathbf{U}\text{sgn}(\mathbf{H}) - \mathbf{U}^*)^\top\|_\infty \\ &\leq \|\mathbf{U}\text{sgn}(\mathbf{H}) - \mathbf{U}^*\|_{2,\infty} \left\{ \|\mathbf{U}^*\|_{2,\infty} + \|\mathbf{U}\text{sgn}(\mathbf{H})\|_{2,\infty} \right\} \\ &\leq \|\mathbf{U}\text{sgn}(\mathbf{H}) - \mathbf{U}^*\|_{2,\infty} \left\{ \|\mathbf{U}^*\|_{2,\infty} + \|\mathbf{U}\|_{2,\infty} \right\}. \end{aligned}$$

Taking this together with the bound (4.31a) in Theorem 4.2, the incoherence assumption, and the inequality (4.133), we arrive at

$$\|\hat{\mathbf{P}} - \mathbf{P}^*\|_\infty \lesssim \frac{(\sigma\kappa\sqrt{\mu r} + \sigma\sqrt{r\log n})}{|\lambda_r^*|} \sqrt{\frac{\mu r}{n}}. \quad (4.174)$$

This taken together with (4.159) indicates that

$$\begin{aligned} \|\hat{\mathbf{P}}\|_\infty &\leq \|\mathbf{P}^*\|_\infty + \|\hat{\mathbf{P}} - \mathbf{P}^*\|_\infty \\ &\lesssim \frac{\mu r}{n} + \frac{(\sigma\kappa\sqrt{\mu r} + \sigma\sqrt{r\log n})}{|\lambda_r^*|} \sqrt{\frac{\mu r}{n}} \asymp \frac{\mu r}{n}, \end{aligned} \quad (4.175)$$

with the proviso that $\sigma\sqrt{n} \lesssim |\lambda_r^*|/\kappa$ and $\sigma\sqrt{n\log n} \lesssim |\lambda_r^*|$.

Armed with the preceding bounds, we are now positioned to control $\hat{v}_{i,j} - \tilde{v}_{i,j}$. From the definition of $\tilde{v}_{i,j}$ and $v_{i,j}^*$, we recognize that

$$\begin{aligned} |\hat{v}_{i,j} - \tilde{v}_{i,j}| &\leq \underbrace{\left| \sum_{l=1}^n \left(\hat{E}_{i,l}^2 \hat{P}_{l,j}^2 - E_{i,l}^2 P_{l,j}^{*2} \right) \right|}_{=: \alpha_1} + \underbrace{\left| \sum_{l=1}^n \left(\hat{P}_{i,l}^2 \hat{E}_{l,j}^2 - P_{i,l}^{*2} E_{l,j}^2 \right) \right|}_{=: \alpha_2} \\ &\quad + 2 \underbrace{\left| \hat{E}_{i,j}^2 \hat{P}_{i,i} \hat{P}_{j,j} - E_{i,j}^2 P_{i,i}^* P_{j,j}^* \right|}_{=: \alpha_3}, \end{aligned} \quad (4.176)$$

leaving us with three terms to cope with. Regarding the first term α_1 on the right-hand side of (4.176), it can be easily verified that

$$\begin{aligned} \alpha_1 &\leq \left| \sum_{l=1}^n \left(\hat{E}_{i,l}^2 \hat{P}_{l,j}^2 - E_{i,l}^2 \hat{P}_{l,j}^2 \right) \right| + \left| \sum_{l=1}^n \left(E_{i,l}^2 \hat{P}_{l,j}^2 - E_{i,l}^2 P_{l,j}^{*2} \right) \right| \\ &\leq \left(\max_l |\hat{E}_{i,l}^2 - E_{i,l}^2| \right) \sum_{l=1}^n \hat{P}_{l,j}^2 + \left(\max_l |\hat{P}_{l,j}^2 - P_{l,j}^{*2}| \right) \sum_{l=1}^n E_{i,l}^2 \\ &\stackrel{(i)}{\leq} \left(\|\mathbf{E}\|_\infty + \|\hat{\mathbf{E}}\|_\infty \right) \|\hat{\mathbf{E}} - \mathbf{E}\|_\infty \hat{P}_{j,j} \\ &\quad + \left(\|\mathbf{P}^*\|_\infty + \|\hat{\mathbf{P}}\|_\infty \right) \|\hat{\mathbf{P}} - \mathbf{P}^*\|_\infty \|\mathbf{E}\|^2 \\ &\stackrel{(ii)}{\lesssim} B\sigma\kappa^2\mu r \sqrt{\frac{\log n}{n}} \cdot \frac{\mu r}{n} + \frac{\mu r}{n} \cdot \frac{(\sigma\kappa\sqrt{\mu r} + \sigma\sqrt{r\log n})}{|\lambda_r^*|} \sqrt{\frac{\mu r}{n}} \cdot \sigma^2 n. \end{aligned}$$

Here, (i) makes use of (4.158) (with \mathbf{P}^* replaced by $\hat{\mathbf{P}}$), whereas (ii) holds true due to (4.172), (4.173), (4.159), (4.174), (4.175), and

Lemma 4.13. The second term α_2 on the right-hand side of (4.176) can be bounded in the same manner and we omit it here for brevity. When it comes to the last term α_3 on the right-hand side of (4.176), one has

$$\begin{aligned} \alpha_3 &\leq \left| \widehat{E}_{i,j}^2 - E_{i,j}^2 \right| \left| \widehat{P}_{i,i} \widehat{P}_{j,j} \right| + E_{i,j}^2 \left| \widehat{P}_{i,i} \widehat{P}_{j,j} - P_{i,i}^* P_{j,j}^* \right| \\ &\lesssim \left\| \widehat{\mathbf{E}} - \mathbf{E} \right\|_\infty \left\{ \left\| \mathbf{E} \right\|_\infty + \left\| \widehat{\mathbf{E}} \right\|_\infty \right\} \left\| \widehat{\mathbf{P}} \right\|_\infty^2 \\ &\quad + B^2 \left\| \widehat{\mathbf{P}} - \mathbf{P}^* \right\|_\infty \left\{ \left\| \mathbf{P}^* \right\|_\infty + \left\| \widehat{\mathbf{P}} \right\|_\infty \right\} \\ &\lesssim \sigma \kappa^2 \mu r \sqrt{\frac{\log n}{n}} \cdot B \cdot \left(\frac{\mu r}{n} \right)^2 + B^2 \cdot \frac{(\sigma \kappa \sqrt{\mu r} + \sigma \sqrt{r \log n})}{|\lambda_r^*|} \sqrt{\frac{\mu r}{n}} \cdot \frac{\mu r}{n} \\ &\lesssim B \sigma \kappa^2 \mu r \sqrt{\frac{\log n}{n}} \cdot \frac{\mu r}{n} + \frac{\mu r}{n} \cdot \frac{(\sigma \kappa \sqrt{\mu r} + \sigma \sqrt{r \log n})}{|\lambda_r^*|} \sqrt{\frac{\mu r}{n}} \cdot \sigma^2 n, \end{aligned}$$

where the penultimate inequality is a consequence of (4.172), (4.173), (4.159), (4.174) and (4.175), and the last line holds true as long as $\mu r \leq n$ (cf. (4.25)) and $B \lesssim \sigma \sqrt{n}$. Combining the above inequalities allows one to reach

$$\begin{aligned} |\widehat{v}_{i,j} - \widetilde{v}_{i,j}| &\leq \alpha_1 + \alpha_2 + \alpha_3 \\ &\lesssim B \sigma \kappa^2 \mu r \sqrt{\frac{\log n}{n}} \cdot \frac{\mu r}{n} + \frac{\mu r}{n} \cdot \frac{\sigma \kappa \sqrt{\mu r} + \sigma \sqrt{r \log n}}{|\lambda_r^*|} \sqrt{\frac{\mu r}{n}} \cdot \sigma^2 n \\ &\asymp \frac{B \sigma \kappa^2 \mu^2 r^2 \sqrt{\log n}}{n^{3/2}} + \frac{\sigma^3 \kappa \mu^2 r^2 + \sigma^3 \mu^{3/2} r^2 \sqrt{\log n}}{|\lambda_r^*| \sqrt{n}} \end{aligned}$$

with probability at least $1 - O(n^{-5})$.

Step 3: combining the above bounds

Putting together the results in the previous steps, we can readily derive

$$\begin{aligned} |\widehat{v}_{i,j} - v_{i,j}^*| &\leq |\widetilde{v}_{i,j} - v_{i,j}^*| + |\widehat{v}_{i,j} - \widetilde{v}_{i,j}| \\ &\lesssim \left(\frac{\sigma B \mu^{3/2} r^{3/2} \sqrt{\log n}}{n^{3/2}} + \frac{\mu^2 r^2 B^2 \log n}{n^2} \right) \\ &\quad + \left(\frac{B \sigma \kappa^2 \mu^2 r^2 \sqrt{\log n}}{n^{3/2}} + \frac{\sigma^3 \kappa \mu^2 r^2 + \sigma^3 \mu^{3/2} r^2 \sqrt{\log n}}{|\lambda_r^*| \sqrt{n}} \right) \\ &\asymp \frac{B \sigma \kappa^2 \mu^2 r^2 \sqrt{\log n}}{n^{3/2}} + \frac{\sigma^3 \kappa \mu^2 r^2 \sqrt{\log n}}{|\lambda_r^*| \sqrt{n}}, \end{aligned} \tag{4.177}$$

where the last relation is guaranteed as long as $B \lesssim \sigma \sqrt{n/\log n}$. Consequently, if Condition (4.167) holds, then it follows from (4.177) and the lower bound (4.162) that

$$|\hat{v}_{i,j} - v_{i,j}^*| \lesssim \frac{1}{\sqrt{\log n}} v_{i,j}^*,$$

thus concluding the proof.

4.13 Notes

Leave-one-out analysis. The core idea of leave-one-out analysis, which drops a small amount of randomness to decouple complicated statistical dependency, is deeply rooted in the probability and statistics literature. For instance, an idea of this kind was invoked by Stein (1972) to help establish normal approximation, was paired with the Stieltjes transform to establish the limiting spectral law of random matrices (see, e.g., (Tao, 2012, Section 2.4.3)), and bears some resemblance to the cavity method in statistical physics (Mezard and Montanari, 2009). When it comes to statistical estimation, a prominent series of work that unveiled the striking effectiveness of leave-one-out analysis was El Karoui *et al.* (2013) and El Karoui (2018), which characterized rigorously the sharp statistical performance (including pre-constants) of M-estimators in high dimension (i.e., a challenging regime where the number of samples is comparable to the number of unknown parameters). The deep analysis framework developed in these papers inspired much of the follow-up work presented in this chapter. Particularly worth mentioning are: (1) Zhong and Boumal (2018): which was the first to determine the entrywise behavior of the generalized projected power method; (2) Abbe *et al.* (2020b) and Chen *et al.* (2019b): which extended the leave-one-out analysis idea to establish entrywise eigenvector perturbation; and (3) Ma *et al.* (2020) and Chen *et al.* (2019a): which characterized tight convergence guarantees for nonconvex optimization algorithms with the aid of leave-one-out ideas. For readers' reference, we list below several topics for which leave-one-out analyses prove useful:

- Maximum likelihood estimation and M-estimation: El Karoui *et al.* (2013), El Karoui (2018), Lei *et al.* (2018), Sur *et al.* (2019), Sur

and Candès (2019), Chen *et al.* (2019b), and Chen *et al.* (2020b);

- spectral methods: Chen *et al.* (2019b), Abbe *et al.* (2020b), Ma *et al.* (2020), Cai *et al.* (2021a), Lei (2019), Abbe *et al.* (2020a), Ling (2020), and Chen *et al.* (2020b);
- nonconvex optimization for statistical estimation: Ma *et al.* (2020), Chen *et al.* (2019a), Li *et al.* (2021c), Chen *et al.* (2020a), Cai *et al.* (2019a), Dong and Shi (2018), Chen *et al.* (2021d), and Wang *et al.* (2021b);
- semidefinite relaxation for low-rank matrix factorization: Zhong and Boumal (2018), Ding and Chen (2020), Chen *et al.* (2020d), Chen *et al.* (2021c), and Chen *et al.* (2021d);
- uncertainty quantification and confidence intervals: Javanmard and Montanari (2018), Chen *et al.* (2019c), Cai *et al.* (2020), and Yan *et al.* (2021);
- cross validation: Xu *et al.* (2019);
- reinforcement learning: Agarwal *et al.* (2020), Li *et al.* (2020a), Pananjady and Wainwright (2020), Zhang *et al.* (2020b), Cui and Yang (2021), and Wang *et al.* (2021a).

ℓ_∞ and $\ell_{2,\infty}$ perturbation theory. The ℓ_∞ and $\ell_{2,\infty}$ perturbation theory for eigenspace and singular subspaces have been investigated in the literature (Fan *et al.*, 2018b; Cape *et al.*, 2019; Eldridge *et al.*, 2018), but only scatteredly until very recently. A modern and systematic framework was established recently, empowered by the leave-one-out analysis idea. Its efficacy and tightness were first demonstrated by Abbe *et al.* (2020b) in a setting that subsumes the one presented herein (with applications to SBMs, phase synchronization and matrix completion), and by Chen *et al.* (2019b) in an asymmetric setting (with application to top- K ranking). The theoretical framework has subsequently been extended in several aspects. For instance, (1) Cai *et al.* (2021a) investigated the “unbalanced” scenario where the column dimension far exceeds the row dimension of the matrix, resulting in

near-optimal fine-grained guarantees for PCA, bi-clustering and tensor completion; (2) Lei (2019) expanded the setting by accounting for more flexible noise distributions (including the ones exhibiting certain dependency structure), leading to tight guarantees for, e.g., spectral clustering with more than two communities, and hierarchical clustering; (3) Abbe *et al.* (2020a) explored a more general ℓ_p perturbation theory that subsumes the ℓ_∞ perturbation theory as special cases. Another plausible approach to study ℓ_∞ eigenvector perturbation is to analyze instead the dynamics of an iterative procedure (e.g., the power method) that converges to the leading eigenvector (Zhong and Boumal, 2018), again using the leave-one-out ideas. This iterative approach offers a perspective complementary to the analysis framework presented herein, while at the same time playing a pivotal role when studying nonconvex optimization algorithms (see Chi *et al.* (2019)). Moving beyond the leave-one-out analysis framework, ℓ_∞ and $\ell_{2,\infty}$ eigenspace perturbation theory has been derived via other powerful tools as well, e.g., the Neumann trick (Eldridge *et al.*, 2018; Chen *et al.*, 2021b; Cheng *et al.*, 2021), the Procrustes analysis (Cape *et al.*, 2019), and more specialized techniques tailored to Gaussian ensembles (Koltchinskii and Xia, 2016; Koltchinskii and Lounici, 2016; Koltchinskii *et al.*, 2020). Perturbations of linear forms and bilinear forms of eigenvectors have also been investigated in the literature (Koltchinskii and Xia, 2016; Koltchinskii and Lounici, 2016; Chen *et al.*, 2021b; Cheng *et al.*, 2021; Fan *et al.*, 2020a; Koltchinskii *et al.*, 2020), which are beyond the scope of this work. Finally, distributional theory and uncertainty quantification for spectral methods, which have been recently studied by Xia (2021), Cheng *et al.* (2021), Fan *et al.* (2020a), Yan *et al.* (2021), and Agterberg *et al.* (2021), are still in their infancy. The results presented in Section 4.6 follow the analysis framework developed in Yan *et al.* (2021).

5

Concluding remarks and open problems

This monograph offered a coherent statistical treatment for spectral methods, resulting in appealing theoretical guarantees for a wide spectrum of data science applications ranging from structured signal reconstruction and factor analysis to clustering and ranking. The important role of statistical thinking cannot be overstated. As has been illuminated, the suite of modern statistical techniques not merely empowers classical ℓ_2 perturbation theory by delivering tight Euclidean error bounds, but also enables fine-grained ℓ_∞ and $\ell_{2,\infty}$ performance guarantees that cannot be derived from classical matrix perturbation theory alone. We highlighted a unified recipe that underlies our application-driven analyses, which will be readily applicable to tackle many other problems.

The vignettes presented herein only reflect the tip of an iceberg regarding the capability of spectral methods. There are multiple aspects about spectral methods that remain inadequately explored and are worthy of future investigation. We conclude this monograph by pointing out a few of them.

- *Precise performance characterization.* The analysis herein falls short of pinpointing a precise trade-off curve between the statistical

accuracy and sample complexity of spectral methods, and might even be off by some logarithmic factor. For algorithms that exhibit order-wise equivalent behavior, comparing their performances requires finer statistical characterization, ideally with sharp pre-constants.

- *Handling dependency structure.* Thus far, the ℓ_∞ and $\ell_{2,\infty}$ perturbation theory we have presented is restricted to the case where the entries of the data samples are independently generated. There is no shortage of applications where the data samples might exhibit across-entry dependency, examples including blind deconvolution (Ahmed *et al.*, 2013) and phase retrieval with coded diffraction patterns (Candès *et al.*, 2015b; Gross *et al.*, 2017). Handling such scenarios might require ideas beyond the current leave-one-out framework.
- *Functional estimation.* In many decision making applications, what ultimately matters might not be full information about an eigenvector of a matrix, but rather, some deterministic functions (e.g., certain linear functionals or polynomials) about the entries of this eigenvector. However, naive “plug-in” estimators—namely, estimating the eigenvector first and plugging it into the target functional—might suffer from significant estimation bias, even in the case of a linear functional. A systematic bias-correction paradigm is therefore needed to enable optimal functional estimation.
- *Small eigengaps.* All theory presented in this monograph imposes a stringent requirement on the associated eigengap, that is, it needs to exceed the spectral norm of a noise or perturbation matrix. While this eigengap criterion might be unavoidable in generic matrix perturbation theory (which takes a worst-case perspective), there is often no statistical lower bound that rules out the possibility of reliable eigenspace estimation when the eigengap drops below the perturbation size. It would be of fundamental importance to understand how a small eigengap impacts the efficacy of spectral methods under various statistical models.

- *Weak and sparse factors.* As mentioned previously, low-rank matrices often admit factor-model interpretations. In many applications, one has to deal with weak factors, on which only a small fraction of the variables have non-negligible loadings. This gives rise to sparse patterns on the loading matrix or the eigenvectors of the covariance matrix. To utilize such a sparsity structure, a simple method is to apply marginal screening techniques (Fan and Lv, 2008; Fan and Fan, 2008). Examples of this kind include supervised PCA (Bair *et al.*, 2006), PCA on “targeted predictors” (Bai and Ng, 2008), and sparse PCA (Zou *et al.*, 2006; Johnstone and Lu, 2009; Ma, 2013). It remains to develop a more systematic and unified theory concerning how to efficiently exploit such special structures in low-rank factorizations, taking into account both statistical and computational considerations.
- *Heterogeneous missing patterns.* When it comes to missing data, the theory presented herein adopts a uniform sampling model where every entry is independently observed with the same probability. In practice, however, one might encounter non-uniform sampling mechanisms, where the sampling probabilities are non-identical across different entries. How to develop an effective spectral method to automatically account for heterogeneous observation patterns, ideally without knowing the detailed sampling probabilities *a priori*?
- *Confidence regions and hypothesis testing for individual eigenvectors.* Given the output of a spectral method, one might be asked to produce valid confidence regions for an unknown individual eigenvector of interest, a task that has not been fully resolved by the existing literature. Another closely related task is hypothesis testing for individual eigenvectors: given two random samples, how to develop viable statistical tests regarding whether these two samples are associated with the same individual eigenvectors or not. An even more challenging task is concerned with performing efficient statistical inference on some deterministic functions of an individual eigenvector, which remains largely unknown.

Acknowledgements

The authors thank the Editor-in-Chief Prof. Michael Jordan for his encouragements, and the publisher Mike Casey for his editorial help.

We are deeply indebted to our wonderful collaborators who have contributed significantly to, and helped shape our perspectives into, the materials presented herein, including Emmanuel Abbe, Changxiao Cai, Emmanuel Candès, Yanxi Chen, Chen Cheng, Yonina Eldar, Yingying Fan, Haoyu Fu, Andrea Goldsmith, Leonidas Guibas, Qixing Huang, Govinda Kamath, Tracy Ke, Gen Li, Yuanxin Li, Yingbin Liang, Yuan Liao, Junwei Lu, Yue Lu, Jinchi Lv, Vincent Monardo, H. Vincent Poor, Changho Suh, Tian Tong, David Tse, Bingyan Wang, Kaizheng Wang, Weichen Wang, Yuting Wei, Yuling Yan, Zhuorang Yang, Huishuai Zhang, Yuchen Zhou, Yiqiao Zhong, and Ziwei Zhu. We owe our particular gratitude to Yuling Yan, who has generously helped with most materials presented in Sections 4.6-4.7. We also thank Bingyan Wang and Chen Dan for their helpful comments about an early version of this monograph, Changxiao Cai for his help in producing Figure 3.5, and Kaizheng Wang for his help in generating Figure 4.2.

We gratefully acknowledge the generous financial support of multiple agencies. More specifically, Y. Chen acknowledges the support by the AFOSR YIP award FA9550-19-1-0030, the ONR grant N00014-19-1-2120, the ARO YIP award W911NF-20-1-0097, the ARO grant W911NF-18-1-0303, the NSF grants CCF-1907661, IIS-1900140, IIS-2100158 and

DMS-2014279, and the Princeton SEAS innovation award; Y. Chi has been supported in part by the ONR under the grants N00014-18-1-2142 and N00014-19-1-2404, by the ARO under the grant W911NF-18-1-0303, and by the NSF under the grants CAREER ECCS-1818571, CCF-1901199, CCF-1806154, CCF-2007911, CCF-2106778 and ECCS-2126634; and J. Fan has been supported in part by the ONR grant N00014-19-1-2120, the NSF grants DMS-1662139, DMS-1712591, DMS-2053832, DMS-2052926, and the NIH grant R01-GM072611, and the Princeton SEAS innovation award. Part of this work was done while Y. Chen was visiting the Simons Institute for the Theory of Computing.

Last but not least, this work would not come to existence without the continuing support of our families, especially during the difficult time of COVID-19 pandemic when this monograph was completed. Y. Chen thanks Yuting Wei for bringing love and encouragement everyday during the writing of this monograph. Y. Chi is deeply grateful to her parents, husband, and daughter for being the silver lining in the pandemic. J. Fan enjoys gratefully his wife and daughters' company and thanks them for compassionate support. C. Ma thanks Xinyi Liu for her unfailing support, and Pidan the Cat for bringing surprises and joys everyday. This monograph is dedicated to them.

References

Abbe, E. (2017). “Community detection and stochastic block models: recent developments”. *Journal of Machine Learning Research*. 18(1): 6446–6531.

Abbe, E., A. Bandeira, and G. Hall. (2016). “Exact recovery in the stochastic block model”. *IEEE Transactions on Information Theory*. 62(1): 471–487.

Abbe, E., J. Fan, and K. Wang. (2020a). “An L_p theory of PCA and spectral clustering”. *arXiv preprint arXiv:2006.14062*.

Abbe, E., J. Fan, K. Wang, and Y. Zhong. (2020b). “Entrywise eigenvector analysis of random matrices with low expected rank”. *The Annals of Statistics*. 48(3): 1452–1474.

Achlioptas, D. and F. McSherry. (2007). “Fast computation of low-rank matrix approximations”. *Journal of the ACM*. 54(2): 9–es.

Agarwal, A., S. Kakade, and L. F. Yang. (2020). “Model-based reinforcement learning with a generative model is minimax optimal”. In: *Conference on Learning Theory*. 67–83.

Agarwal, A., P. Patil, and S. Agarwal. (2018). “Accelerated spectral ranking”. In: *International Conference on Machine Learning*. 70–79.

Agterberg, J., Z. Lubberts, and C. Priebe. (2021). “Entrywise estimation of singular vectors of low-rank matrices with heteroskedasticity and dependence”. *arXiv preprint arXiv:2105.13346*.

Ahmed, A., B. Recht, and J. Romberg. (2013). “Blind deconvolution using convex programming”. *IEEE Transactions on Information Theory*. 60(3): 1711–1732.

Ahn, K., K. Lee, and C. Suh. (2018). “Hypergraph spectral clustering in the weighted stochastic block model”. *IEEE Journal of Selected Topics in Signal Processing*. 12(5): 959–974.

Ahn, S. C. and A. R. Horenstein. (2013). “Eigenvalue ratio test for the number of factors”. *Econometrica*. 81(3): 1203–1227.

Aloise, D., A. Deshpande, P. Hansen, and P. Popat. (2009). “NP-hardness of Euclidean sum-of-squares clustering”. *Machine learning*. 75(2): 245–248.

Alon, N., M. Krivelevich, and B. Sudakov. (1998). “Finding a large hidden clique in a random graph”. *Random Structures & Algorithms*. 13(3-4): 457–466.

Amini, A. A. and M. J. Wainwright. (2008). “High-dimensional analysis of semidefinite relaxations for sparse principal components”. In: *International Symposium on Information Theory*. 2454–2458.

Amini, A. A., A. Chen, P. J. Bickel, and E. Levina. (2013). “Pseudo-likelihood methods for community detection in large sparse networks”. *The Annals of Statistics*. 41(4): 2097–2122.

Amjad, M., D. Shah, and D. Shen. (2018). “Robust synthetic control”. *The Journal of Machine Learning Research*. 19(1): 802–852.

Anandkumar, A., R. Ge, D. Hsu, S. M. Kakade, and M. Telgarsky. (2014). “Tensor decompositions for learning latent variable models”. *Journal of Machine Learning Research*. 15: 2773–2832.

Anderson, T. W. (1962). *An introduction to multivariate statistical analysis*. Wiley New York.

Awasthi, P., A. S. Bandeira, M. Charikar, R. Krishnaswamy, S. Villar, and R. Ward. (2015). “Relax, no need to round: Integrality of clustering formulations”. In: *Conference on Innovations in Theoretical Computer Science*. 191–200.

Awasthi, P. and O. Sheffet. (2012). “Improved spectral-norm bounds for clustering”. In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*. Springer. 37–49.

Azaouzi, M., D. Rhouma, and L. B. Romdhane. (2019). “Community detection in large-scale social networks: state-of-the-art and future directions”. *Social Network Analysis and Mining*. 9(1): 23.

Bahmani, S. and J. Romberg. (2017). “Phase retrieval meets statistical learning theory: A flexible convex relaxation”. In: *Artificial Intelligence and Statistics*. 252–260.

Bai, J. (2003). “Inferential theory for factor models of large dimensions”. *Econometrica*. 71(1): 135–171.

Bai, J. (2009). “Panel data models with interactive fixed effects”. *Econometrica*. 77(4): 1229–1279.

Bai, J. and S. Ng. (2002). “Determining the number of factors in approximate factor models”. *Econometrica*. 70(1): 191–221.

Bai, J. and S. Ng. (2008). “Forecasting economic time series using targeted predictors”. *Journal of Econometrics*. 146(2): 304–317.

Bai, J. and P. Wang. (2016). “Econometric analysis of large factor models”. *Annual Review of Economics*. 8: 53–80.

Bair, E., T. Hastie, D. Paul, and R. Tibshirani. (2006). “Prediction by supervised principal components”. *Journal of the American Statistical Association*. 101(473): 119–137.

Bajaj, C., T. Gao, Z. He, Q. Huang, and Z. Liang. (2018). “SMAC: Simultaneous mapping and clustering using spectral decompositions”. In: *International Conference on Machine Learning*. 324–333.

Balakrishnan, S., M. J. Wainwright, and B. Yu. (2017). “Statistical guarantees for the EM algorithm: From population to sample-based analysis”. *The Annals of Statistics*. 45(1): 77–120.

Balakrishnan, S., M. Xu, A. Krishnamurthy, and A. Singh. (2011). “Noise thresholds for spectral clustering”. In: *Advances in Neural Information Processing Systems*. 954–962.

Balzano, L., Y. Chi, and Y. M. Lu. (2018). “Streaming PCA and subspace tracking: The missing data case”. *Proceedings of the IEEE*. 106(8): 1293–1310.

Bandeira, A. S., N. Boumal, and A. Singer. (2017). “Tightness of the maximum likelihood semidefinite relaxation for angular synchronization”. *Mathematical Programming*. 163(1-2): 145–167.

Bandeira, A. S. and R. Van Handel. (2016). “Sharp nonasymptotic bounds on the norm of random matrices with independent entries”. *The Annals of Probability*. 44(4): 2479–2506.

Bao, Z., X. Ding, J. Wang, and K. Wang. (2020). “Statistical inference for principal components of spiked covariance matrix”. *arXiv preprint arXiv:2008.11903*.

Bao, Z., X. Ding, and K. Wang. (2021). “Singular vector and singular subspace distribution for the matrix denoising model”. *The Annals of Statistics*. 49(1): 370–392.

Barak, B. and A. Moitra. (2016). “Noisy tensor completion via the sum-of-squares hierarchy”. In: *Conference on Learning Theory*. PMLR. 417–445.

Belkin, M. and P. Niyogi. (2003). “Laplacian eigenmaps for dimensionality reduction and data representation”. *Neural Computation*. 15(6): 1373–1396.

Bhatia, R. (2013). *Matrix analysis*. Vol. 169. Springer Science & Business Media.

Binkiewicz, N., J. T. Vogelstein, and K. Rohe. (2017). “Covariate-assisted spectral clustering”. *Biometrika*. 104(2): 361–377.

Boucheron, S., G. Lugosi, and P. Massart. (2013). *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press.

Boumal, N. and P.-A. Absil. (2015). “Low-rank matrix completion via preconditioned optimization on the Grassmann manifold”. *Linear Algebra and its Applications*. 475: 200–239.

Bradley, R. A. and M. E. Terry. (1952). “Rank analysis of incomplete block designs: I. The method of paired comparisons”. *Biometrika*. 39(3/4): 324–345.

Brémaud, P. (2013). *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*. Vol. 31. Springer Science & Business Media.

Browet, A., P.-A. Absil, and P. Van Dooren. (2011). “Community detection for hierarchical image segmentation”. In: *International Workshop on Combinatorial Image Analysis*. 358–371.

Cai, C., G. Li, Y. Chi, H. V. Poor, and Y. Chen. (2021a). “Subspace estimation from unbalanced and incomplete data matrices: $\ell_{2,\infty}$ statistical guarantees”. *The Annals of Statistics*. 49(2): 944–967.

Cai, C., G. Li, H. V. Poor, and Y. Chen. (2019a). “Nonconvex low-rank symmetric tensor completion from noisy data”. In: *Advances in Neural Information Processing Systems*. 1863–1874.

Cai, C., G. Li, H. V. Poor, and Y. Chen. (2021b). “Nonconvex low-rank tensor completion from noisy data”. *Operations Research*.

Cai, C., H. V. Poor, and Y. Chen. (2020). “Uncertainty quantification for nonconvex tensor completion: Confidence intervals, heteroscedasticity and optimality”. In: *International Conference on Machine Learning*. 1271–1282.

Cai, J.-F., H. Liu, and Y. Wang. (2019b). “Fast rank-one alternating minimization algorithm for phase retrieval”. *Journal of Scientific Computing*. 79(1): 128–147.

Cai, T. T. and X. Li. (2015). “Robust and computationally feasible community detection in the presence of arbitrary outlier nodes”. *The Annals of Statistics*. 43(3): 1027–1059.

Cai, T. T., X. Li, and Z. Ma. (2016). “Optimal rates of convergence for noisy sparse phase retrieval via thresholded Wirtinger flow”. *The Annals of Statistics*. 44(5): 2221–2251.

Cai, T. T., Z. Ma, and Y. Wu. (2013). “Sparse PCA: Optimal rates and adaptive estimation”. *The Annals of Statistics*. 41(6): 3074–3110.

Cai, T. T. and A. Zhang. (2015). “ROP: Matrix recovery via rank-one projections”. *The Annals of Statistics*. 43(1): 102–138.

Cai, T. T. and A. Zhang. (2018). “Rate-optimal perturbation bounds for singular subspaces with applications to high-dimensional statistics”. *The Annals of Statistics*. 46(1): 60–89.

Cambareri, V. and L. Jacques. (2016). “A non-convex blind calibration method for randomised sensing strategies”. In: *2016 4th International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa)*. 16–20.

Candès, E., X. Li, and M. Soltanolkotabi. (2015a). “Phase retrieval via Wirtinger flow: Theory and algorithms”. *IEEE Transactions on Information Theory*. 61(4): 1985–2007.

Candès, E. J. (2014). “Mathematics of sparsity (and a few other things)”. In: *Proceedings of the International Congress of Mathematicians, Seoul, South Korea*. Vol. 123.

Candès, E. J., X. Li, Y. Ma, and J. Wright. (2011). “Robust principal component analysis?” *Journal of the ACM (JACM)*. 58(3): 1–37.

Candès, E. J., X. Li, and M. Soltanolkotabi. (2015b). “Phase retrieval from coded diffraction patterns”. *Applied and Computational Harmonic Analysis*. 39(2): 277–299.

Candès, E. J. and Y. Plan. (2010). “Matrix completion with noise”. *Proceedings of the IEEE*. 98(6): 925–936.

Candès, E. J., T. Strohmer, and V. Voroninski. (2013). “Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming”. *Communications on Pure and Applied Mathematics*. 66(8): 1241–1274.

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

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

Cao, Y. and Y. Xie. (2015). “Poisson matrix recovery and completion”. *IEEE Transactions on Signal Processing*. 64(6): 1609–1620.

Cape, J., M. Tang, and C. E. Priebe. (2019). “The two-to-infinity norm and singular subspace geometry with applications to high-dimensional statistics”. *The Annals of Statistics*. 47(5): 2405–2439.

Caplin, A. and B. Nalebuff. (1991). “Aggregation and social choice: A mean voter theorem”. *Econometrica: Journal of the Econometric Society*: 1–23.

Chandra, R., T. Goldstein, and C. Studer. (2019). “Phasepack: A phase retrieval library”. In: *2019 13th International conference on Sampling Theory and Applications (SampTA)*. IEEE. 1–5.

Chandrasekaran, V., P. A. Parrilo, and A. S. Willsky. (2010). “Latent variable graphical model selection via convex optimization”. In: *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE. 1610–1613.

Chandrasekaran, V., S. Sanghavi, P. A. Parrilo, and A. S. Willsky. (2011). “Rank-sparsity incoherence for matrix decomposition”. *SIAM Journal on Optimization*. 21(2): 572–596.

Charisopoulos, V., Y. Chen, D. Davis, M. Diaz, L. Ding, and D. Drusvyatskiy. (2021). “Low-rank matrix recovery with composite optimization: good conditioning and rapid convergence”. *Foundations of Computational Mathematics*: 1–89.

Charisopoulos, V., D. Davis, M. Diaz, and D. Drusvyatskiy. (2019). “Composite optimization for robust blind deconvolution”. *arXiv preprint arXiv:1901.01624*.

Chatterjee, S. (2014). “Matrix estimation by universal singular value thresholding”. *The Annals of Statistics*. 43(1): 177–214.

Chaudhuri, K., F. Chung, and A. Tsiatas. (2012). “Spectral clustering of graphs with general degrees in the extended planted partition model”. In: *Conference on Learning Theory*. 35.1–35.23.

Chen, P.-H. C., J. Chen, Y. Yeshurun, U. Hasson, J. Haxby, and P. J. Ramadge. (2015a). “A reduced-dimension fMRI shared response model”. In: *Advances in Neural Information Processing Systems*. 460–468.

Chen, J., D. Liu, and X. Li. (2020a). “Nonconvex rectangular matrix completion via gradient descent without $\ell_{2,\infty}$ regularization”. *IEEE Transactions on Information Theory*. 66(9): 5806–5841.

Chen, L. H., L. Goldstein, and Q.-M. Shao. (2010). *Normal approximation by Stein’s method*. Springer Science & Business Media.

Chen, P., A. Fannjiang, and G.-R. Liu. (2017). “Phase retrieval by linear algebra”. *SIAM Journal on Matrix Analysis and Applications*. 38(3): 854–868.

Chen, P.-Y. and A. O. Hero. (2015). “Phase transitions in spectral community detection”. *IEEE Transactions on Signal Processing*. 63(16): 4339–4347.

Chen, P., C. Gao, and A. Y. Zhang. (2020b). “Partial recovery for top- K ranking: Optimality of MLE and sub-optimality of spectral method”. *arXiv preprint arXiv:2006.16485*.

Chen, S., S. Liu, and Z. Ma. (2020c). “Global and individualized community detection in inhomogeneous multilayer networks”. *arXiv preprint arXiv:2012.00933*.

Chen, X., P. N. Bennett, K. Collins-Thompson, and E. Horvitz. (2013). “Pairwise ranking aggregation in a crowdsourced setting”. In: *International conference on Web search and data mining*. 193–202.

Chen, X. and Y. Yang. (2020). “Cutoff for exact recovery of Gaussian mixture models”. *arXiv preprint arXiv:2001.01194*.

Chen, Y., Y. Chi, and A. Goldsmith. (2015b). “Exact and stable covariance estimation from quadratic sampling via convex programming”. *IEEE Transactions on Information Theory*. 61(7): 4034–4059.

Chen, Y., C. Ma, H. V. Poor, and Y. Chen. (2021a). “Learning mixtures of low-rank models”. *IEEE Transactions on Information Theory*. 67(7): 4613–4636.

Chen, Y. and Y. Chi. (2018). “Harnessing structures in big data via guaranteed low-rank matrix estimation: Recent theory and fast algorithms via convex and nonconvex optimization”. *IEEE Signal Processing Magazine*. 35(4): 14–31.

Chen, Y. and M. J. Wainwright. (2015). “Fast low-rank estimation by projected gradient descent: General statistical and algorithmic guarantees”. *arXiv preprint arXiv:1509.03025*.

Chen, Y. and E. Candès. (2017). “Solving random quadratic systems of equations is nearly as easy as solving linear systems”. *Communications on Pure and Applied Mathematics*. 70(5): 822–883.

Chen, Y. and E. Candès. (2018). “The projected power method: An efficient algorithm for joint alignment from pairwise differences”. *Communications on Pure and Applied Mathematics*. 71(8): 1648–1714.

Chen, Y., C. Cheng, and J. Fan. (2021b). “Asymmetry helps: Eigenvalue and eigenvector analyses of asymmetrically perturbed low-rank matrices”. *The Annals of Statistics*. 49(1): 435–458.

Chen, Y. and Y. Chi. (2014). “Robust spectral compressed sensing via structured matrix completion”. *IEEE Transactions on Information Theory*. 60(10): 6576–6601.

Chen, Y., Y. Chi, J. Fan, and C. Ma. (2019a). “Gradient descent with random initialization: Fast global convergence for nonconvex phase retrieval”. *Mathematical Programming*. 176(1-2): 5–37.

Chen, Y., Y. Chi, J. Fan, C. Ma, and Y. Yan. (2020d). “Noisy matrix completion: Understanding statistical guarantees for convex relaxation via nonconvex optimization”. *SIAM Journal on Optimization*. 30(4): 3098–3121.

Chen, Y., J. Fan, C. Ma, and K. Wang. (2019b). “Spectral method and regularized MLE are both optimal for top- K ranking”. *The Annals of Statistics*. 47(4): 2204–2235.

Chen, Y., J. Fan, C. Ma, and Y. Yan. (2019c). “Inference and uncertainty quantification for noisy matrix completion”. *Proceedings of the National Academy of Sciences*. 116(46): 22931–22937.

Chen, Y., J. Fan, C. Ma, and Y. Yan. (2021c). “Bridging convex and nonconvex optimization in robust PCA: Noise, outliers, and missing data”. *accepted to The Annals of Statistics*.

Chen, Y., J. Fan, B. Wang, and Y. Yan. (2021d). “Convex and nonconvex optimization are both minimax-optimal for noisy blind deconvolution”. *accepted to Journal of the American Statistical Association*.

Chen, Y., L. J. Guibas, and Q.-X. Huang. (2014). “Near-optimal joint object matching via convex relaxation”. In: *International Conference on Machine Learning (ICML)*. 100–108.

Chen, Y., G. Kamath, C. Suh, and D. Tse. (2016a). “Community recovery in graphs with locality”. In: *International Conference on Machine Learning*. 689–698.

Chen, Y. and C. Suh. (2015). “Spectral MLE: Top- K rank aggregation from pairwise comparisons”. In: *International Conference on Machine Learning*. 371–380.

Chen, Y., C. Suh, and A. J. Goldsmith. (2016b). “Information recovery from pairwise measurements”. *IEEE Transactions on Information Theory*. 62(10): 5881–5905.

Cheng, C., Y. Wei, and Y. Chen. (2021). “Tackling small eigen-gaps: Fine-grained eigenvector estimation and inference under heteroscedastic noise”. *accepted to IEEE Transactions on Information Theory*.

Cherapanamjeri, Y., K. Gupta, and P. Jain. (2017). “Nearly optimal robust matrix completion”. In: *International Conference on Machine Learning*. 797–805.

Chernozhukov, V., C. Hansen, Y. Liao, and Y. Zhu. (2021). “Inference for low-rank models”. *arXiv preprint arXiv:2107.02602*.

Chi, Y. and H. Fu. (2017). “Subspace learning from bits”. *IEEE Transactions on Signal Processing*. 65(17): 4429–4442.

Chi, Y., Y. M. Lu, and Y. Chen. (2019). “Nonconvex optimization meets low-rank matrix factorization: An overview”. *IEEE Transactions on Signal Processing*. 67(20): 5239–5269.

Chin, P., A. Rao, and V. Vu. (2015). “Stochastic block model and community detection in sparse graphs: A spectral algorithm with optimal rate of recovery”. In: *Conference on Learning Theory*. 391–423.

Coja-Oghlan, A. (2010). “Graph partitioning via adaptive spectral techniques”. *Combinatorics, Probability & Computing*. 19(2): 227.

Cole, S. and Y. Zhu. (2020). “Exact recovery in the hypergraph stochastic block model: A spectral algorithm”. *Linear Algebra and its Applications*. 593: 45–73.

Cook, R. D. (2007). “Fisher lecture: Dimension reduction in regression”. *Statistical Science*. 22(1): 1–26.

Cui, Q. and L. F. Yang. (2021). “Minimax Sample Complexity for Turn-based Stochastic Game”. In: *Conference on Uncertainty in Artificial Intelligence*.

Dalvi, N., A. Dasgupta, R. Kumar, and V. Rastogi. (2013). “Aggregating crowdsourced binary ratings”. In: *International conference on World Wide Web*. 285–294.

Dan, C., Y. Wei, and P. Ravikumar. (2020). “Sharp Statistical Guarantees for Adversarially Robust Gaussian Classification”. In: *International Conference on Machine Learning*. 2345–2355.

Dasgupta, S. (1999). “Learning mixtures of Gaussians”. In: *40th Annual Symposium on Foundations of Computer Science*. IEEE. 634–644.

Davenport, M. A., Y. Plan, E. Van Den Berg, and M. Wootters. (2014). “1-bit matrix completion”. *Information and Inference: A Journal of the IMA*. 3(3): 189–223.

Davenport, M. A. and J. Romberg. (2016). “An overview of low-rank matrix recovery from incomplete observations”. *IEEE Journal of Selected Topics in Signal Processing*. 10(4): 608–622.

Davis, C. and W. M. Kahan. (1970). “The rotation of eigenvectors by a perturbation. III”. *SIAM Journal on Numerical Analysis*. 7(1): 1–46.

Dhifallah, O., C. Thrampoulidis, and Y. M. Lu. (2017). “Phase retrieval via linear programming: Fundamental limits and algorithmic improvements”. In: *2017 55th Annual Allerton Conference on Communication, Control, and Computing*. IEEE. 1071–1077.

Diaconis, P. and L. Saloff-Coste. (1993). “Comparison theorems for reversible Markov chains”. *The Annals of Applied Probability*: 696–730.

Ding, L. and Y. Chen. (2020). “Leave-one-out approach for matrix completion: Primal and dual analysis”. *IEEE Transactions on Information Theory*. 66(11): 7274–7301.

Ding, X. (2020). “High dimensional deformed rectangular matrices with applications in matrix denoising”. *Bernoulli*. 26(1): 387–417.

Dong, J. and Y. Shi. (2018). “Nonconvex demixing from bilinear measurements”. *IEEE Transactions on Signal Processing*. 66(19): 5152–5166.

Dopico, F. M. (2000). “A note on $\sin \Theta$ theorems for singular subspace variations”. *BIT Numerical Mathematics*. 40(2): 395–403.

Du, S. S., W. Hu, S. M. Kakade, J. D. Lee, and Q. Lei. (2020). “Few-shot learning via learning the representation, provably”. In: *International Conference on Learning Representations*.

Duan, N. and K.-C. Li. (1991). “Slicing regression: a link-free regression method”. *The Annals of Statistics*: 505–530.

Duan, Y., T. Ke, and M. Wang. (2019). “State aggregation learning from Markov transition data”. In: *Advances in Neural Information Processing Systems*. 4486–4495.

Duchi, J. C. and F. Ruan. (2019). “Solving (most) of a set of quadratic equalities: Composite optimization for robust phase retrieval”. *Information and Inference: A Journal of the IMA*. 8(3): 471–529.

Dudeja, R., M. Bakhshizadeh, J. Ma, and A. Maleki. (2020). “Analysis of spectral methods for phase retrieval with random orthogonal matrices”. *IEEE Transactions on Information Theory*. 66(8): 5182–5203.

Durrett, R. (2007). *Random graph dynamics*. Vol. 200. No. 7.

Dwork, C., R. Kumar, M. Naor, and D. Sivakumar. (2001). “Rank aggregation methods for the web”. In: *International conference on World Wide Web*. 613–622.

El Karoui, N. (2018). “On the impact of predictor geometry on the performance on high-dimensional ridge-regularized generalized robust regression estimators”. *Probability Theory and Related Fields*. 170(1-2): 95–175.

El Karoui, N., D. Bean, P. J. Bickel, C. Lim, and B. Yu. (2013). “On robust regression with high-dimensional predictors”. *Proceedings of the National Academy of Sciences*. 110(36): 14557–14562.

Eldar, Y. C. and S. Mendelson. (2014). “Phase retrieval: Stability and recovery guarantees”. *Applied and Computational Harmonic Analysis*. 36(3): 473–494.

Eldridge, J., M. Belkin, and Y. Wang. (2018). “Unperturbed: Spectral analysis beyond Davis-Kahan”. In: *Algorithmic Learning Theory*. 321–358.

Eriksson, B., L. Balzano, and R. Nowak. (2012). “High-rank matrix completion”. In: *Artificial Intelligence and Statistics*. 373–381.

Fan, J. and Y. Fan. (2008). “High dimensional classification using features annealed independence rules”. *Annals of statistics*. 36(6): 2605.

Fan, J., Y. Fan, X. Han, and J. Lv. (2019a). “SIMPLE: Statistical inference on membership profiles in large networks”. *arXiv preprint arXiv:1910.01734*.

Fan, J., Y. Fan, X. Han, and J. Lv. (2020a). “Asymptotic theory of eigenvectors for random matrices with diverging spikes”. *Journal of the American Statistical Association*: 1–63.

Fan, J., F. Han, and H. Liu. (2014). “Challenges of big data analysis”. *National science review*. 1(2): 293–314.

Fan, J., K. Li, and Y. Liao. (2021a). “Recent developments on factor models and applications in econometric learning”. *Annual Review of Financial Economics*. 13: 1–35.

Fan, J., R. Li, C.-H. Zhang, and H. Zou. (2020b). *Statistical foundations of data science*. CRC press.

Fan, J., Y. Liao, and M. Mincheva. (2013). “Large covariance estimation by thresholding principal orthogonal complements”. *Journal of the Royal Statistical Society. Series B, Statistical methodology*. 75(4): 603–680.

Fan, J., Y. Liao, and W. Wang. (2016). “Projected principal component analysis in factor models”. *The Annals of Statistics*. 44(1): 219–254.

Fan, J., Y. Liao, and J. Yao. (2015). “Power enhancement in high-dimensional cross-sectional tests”. *Econometrica*. 83(4): 1497–1541.

Fan, J. and J. Lv. (2008). “Sure independence screening for ultrahigh dimensional feature space”. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 70(5): 849–911.

Fan, J., Q. Sun, W.-X. Zhou, and Z. Zhu. (2018a). “Principal component analysis for big data”. In: *Wiley StatsRef: Statistics Reference Online*. Wiley Online Library. 1–13.

Fan, J., D. Wang, K. Wang, and Z. Zhu. (2019b). “Distributed estimation of principal eigenspaces”. *The Annals of Statistics*. 47(6): 3009–3031.

Fan, J., K. Wang, Y. Zhong, and Z. Zhu. (2021b). “Robust high-dimensional factor models with applications to statistical machine learning”. *Statistical Science*. 36(2): 303–327.

Fan, J., W. Wang, and Y. Zhong. (2018b). “An ℓ_∞ eigenvector perturbation bound and its application to robust covariance estimation”. *Journal of Machine Learning Research*. 18: 1–42.

Fan, Z., I. M. Johnstone, and Y. Sun. (2018c). “Spiked covariances and principal components analysis in high-dimensional random effects models”. *arXiv preprint arXiv:1806.09529*.

Fannjiang, A. and T. Strohmer. (2020). “The numerics of phase retrieval”. *Acta Numerica*. 29: 125–228.

Fei, Y. and Y. Chen. (2018a). “Exponential error rates of SDP for block models: Beyond Grothendieck’s inequality”. *IEEE Transactions on Information Theory*. 65(1): 551–571.

Fei, Y. and Y. Chen. (2018b). “Hidden integrality of SDP relaxations for sub-Gaussian mixture models”. In: *Conference On Learning Theory*. PMLR. 1931–1965.

Fei, Y. and Y. Chen. (2019). “Achieving the Bayes error rate in stochastic block model by SDP, robustly”. In: *Conference on Learning Theory*. 1235–1269.

Ferguson, T. S. (1996). *A course in large sample theory*. Vol. 38. Chapman & Hall/CRC.

Fienup, J. R. (1982). “Phase retrieval algorithms: A comparison.” *Applied optics*. 21(15): 2758–2769.

Fishkind, D. E., D. L. Sussman, M. Tang, J. T. Vogelstein, and C. E. Priebe. (2013). “Consistent adjacency-spectral partitioning for the stochastic block model when the model parameters are unknown”. *SIAM Journal on Matrix Analysis and Applications*. 34(1): 23–39.

Florescu, L. and W. Perkins. (2016). “Spectral thresholds in the bipartite stochastic block model”. In: *Conference on Learning Theory*. 943–959.

Ford Jr, L. R. (1957). “Solution of a ranking problem from binary comparisons”. *The American Mathematical Monthly*. 64(8P2): 28–33.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin. (2000). “The generalized dynamic-factor model: Identification and estimation”. *Review of Economics and statistics*. 82(4): 540–554.

Fortunato, S. and D. Hric. (2016). “Community detection in networks: A user guide”. *Physics Reports*. 659: 1–44.

Foucart, S., D. Needell, R. Pathak, Y. Plan, and M. Wootters. (2020). “Weighted matrix completion from non-random, non-uniform sampling patterns”. *IEEE Transactions on Information Theory*. 67(2): 1264–1290.

Fu, H., Y. Chi, and Y. Liang. (2020). “Guaranteed recovery of one-hidden-layer neural networks via cross entropy”. *IEEE Transactions on Signal Processing*. 68: 3225–3235.

Gao, C., Z. Ma, A. Y. Zhang, and H. H. Zhou. (2017). “Achieving optimal misclassification proportion in stochastic block models”. *Journal of Machine Learning Research*. 18(1): 1980–2024.

Ge, R., C. Jin, P. Netrapalli, and A. Sidford. (2016). “Efficient algorithms for large-scale generalized eigenvector computation and canonical correlation analysis”. In: *International Conference on Machine Learning*. 2741–2750.

Georghiades, A. S., P. N. Belhumeur, and D. J. Kriegman. (2001). “From few to many: Illumination cone models for face recognition under variable lighting and pose”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 23(6): 643–660.

Ghosh, A., S. Kale, and P. McAfee. (2011). “Who moderates the moderators? crowdsourcing abuse detection in user-generated content”. In: *ACM conference on Electronic commerce*. 167–176.

Ghosh, A. and R. Kannan. (2020). “Alternating minimization converges super-linearly for mixed linear regression”. In: *International Conference on Artificial Intelligence and Statistics*. PMLR. 1093–1103.

Giraud, C. and N. Verzelen. (2019). “Partial recovery bounds for clustering with the relaxed K -means”. *Mathematical Statistics and Learning*. 1(3): 317–374.

Goldstein, T. and C. Studer. (2018). “Phasemax: Convex phase retrieval via basis pursuit”. *IEEE Transactions on Information Theory*. 64(4): 2675–2689.

Gross, D. (2011). “Recovering low-rank matrices from few coefficients in any basis”. *IEEE Transactions on Information Theory*. 57(3): 1548–1566.

Gross, D., F. Krahmer, and R. Kueng. (2017). “Improved recovery guarantees for phase retrieval from coded diffraction patterns”. *Applied and Computational Harmonic Analysis*. 42(1): 37–64.

Hajek, B., S. Oh, and J. Xu. (2014). “Minimax-optimal inference from partial rankings”. *Advances in Neural Information Processing Systems*. 27: 1475–1483.

Hajek, B., Y. Wu, and J. Xu. (2016). “Achieving exact cluster recovery threshold via semidefinite programming”. *IEEE Transactions on Information Theory*. 62(5): 2788–2797.

Han, R., R. Willett, and A. Zhang. (2020). “An optimal statistical and computational framework for generalized tensor estimation”. *arXiv preprint arXiv:2002.11255*.

Han, X., X. Tong, and Y. Fan. (2021). “Eigen selection in spectral clustering: a theory guided practice”. *Journal of the American Statistical Association*. (just-accepted): 1–33.

Han, X., Q. Yang, and Y. Fan. (2019). “Universal rank inference via residual subsampling with application to large networks”. *arXiv preprint arXiv:1912.11583*.

Hand, P. (2017). “Phaselift is robust to a constant fraction of arbitrary errors”. *Applied and Computational Harmonic Analysis*. 42(3): 550–562.

Hand, P. and V. Voroninski. (2016). “Corruption robust phase retrieval via linear programming”. *arXiv preprint arXiv:1612.03547*.

Hansen, L. P. (1982). “Large sample properties of generalized method of moments estimators”. *Econometrica: Journal of the Econometric Society*: 1029–1054.

Hardt, M. (2014). “Understanding alternating minimization for matrix completion”. In: *Foundations of Computer Science (FOCS), 2014 IEEE 55th Annual Symposium on*. IEEE. 651–660.

Hillar, C. J. and L.-H. Lim. (2013). “Most tensor problems are NP-hard”. *Journal of the ACM*. 60(6): 1–39.

Holland, P. W., K. B. Laskey, and S. Leinhardt. (1983). “Stochastic blockmodels: First steps”. *Social networks*. 5(2): 109–137.

Hopkins, S. B., T. Schramm, J. Shi, and D. Steurer. (2016). “Fast spectral algorithms from sum-of-squares proofs: Tensor decomposition and planted sparse vectors”. In: *Symposium on Theory of Computing*. 178–191.

Horn, R. A. and C. R. Johnson. (2012). *Matrix analysis*. Cambridge university press.

Howard, S. R., A. Ramdas, J. McAuliffe, and J. Sekhon. (2020). “Time-uniform Chernoff bounds via nonnegative supermartingales”. *Probability Surveys*. 17: 257–317.

Hsu, D. (2016). “COMS 4772: advanced machine learning”. *Lecture notes, Columbia University*.

Hsu, D. and S. M. Kakade. (2013). “Learning mixtures of spherical Gaussians: Moment methods and spectral decompositions”. In: *Conference on Innovations in Theoretical Computer Science*. 11–20.

Huang, Q., Z. Liang, H. Wang, S. Zuo, and C. Bajaj. (2019a). “Tensor maps for synchronizing heterogeneous shape collections”. *ACM Transactions on Graphics (TOG)*. 38(4): 1–18.

Huang, W. and P. Hand. (2018). “Blind deconvolution by a steepest descent algorithm on a quotient manifold”. *SIAM Journal on Imaging Sciences*. 11(4): 2757–2785.

Huang, X., Z. Liang, X. Zhou, Y. Xie, L. J. Guibas, and Q. Huang. (2019b). “Learning transformation synchronization”. In: *IEEE Conference on Computer Vision and Pattern Recognition*. 8082–8091.

Huang, Q.-X. and L. Guibas. (2013). “Consistent shape maps via semidefinite programming”. In: *Computer Graphics Forum*. Vol. 32. No. 5. 177–186.

Huber, P. J. (2004). *Robust statistics*. Vol. 523. John Wiley & Sons.

Hunter, D. R. (2004). “MM algorithms for generalized Bradley-Terry models”. *The annals of statistics*. 32(1): 384–406.

Iguchi, T., D. G. Mixon, J. Peterson, and S. Villar. (2017). “Probably certifiably correct k -means clustering”. *Mathematical Programming*. 165(2): 605–642.

Jaganathan, K., Y. C. Eldar, and B. Hassibi. (2016). “Phase retrieval: An overview of recent developments”. *Optical Compressive Imaging*: 263–296.

Jagatap, G. and C. Hegde. (2019). “Sample-efficient algorithms for recovering structured signals from magnitude-only measurements”. *IEEE Transactions on Information Theory*. 65(7): 4434–4456.

Jain, P. and P. Kar. (2017). “Non-convex optimization for machine learning”. *Foundations and Trends® in Machine Learning*. 10(3-4): 142–363.

Jain, P., P. Netrapalli, and S. Sanghavi. (2013). “Low-rank matrix completion using alternating minimization”. In: *Symposium on Theory of Computing*. 665–674.

Jalali, A., Y. Chen, S. Sanghavi, and H. Xu. (2011). “Clustering partially observed graphs via convex optimization”. In: *International Conference on Machine Learning*. 1001–1008.

Jang, M., S. Kim, C. Suh, and S. Oh. (2016). “Top- K ranking from pairwise comparisons: When spectral ranking is optimal”. *arXiv preprint arXiv:1603.04153*.

Janzamin, M., R. Ge, J. Kossaifi, and A. Anandkumar. (2019). “Spectral learning on matrices and tensors”. *Foundations and Trends in Machine Learning*. 12(5-6): 393–536.

Javanmard, A. and A. Montanari. (2013). “Localization from incomplete noisy distance measurements”. *Foundations of Computational Mathematics*. 13(3): 297–345.

Javanmard, A. and A. Montanari. (2018). “Debiasing the Lasso: Optimal sample size for Gaussian designs”. *The Annals of Statistics*. 46(6A): 2593–2622.

Jeong, H. and C. S. Güntürk. (2017). “Convergence of the randomized Kaczmarz method for phase retrieval”. *arXiv preprint arXiv:1706.10291*.

Jiang, S., Y. Liu, and Y. Tang. (2017). “A unified perturbation analysis framework for countable Markov chains”. *Linear Algebra and its Applications*. 529: 413–440.

Jin, C., S. M. Kakade, and P. Netrapalli. (2016a). “Provable efficient online matrix completion via non-convex stochastic gradient descent”. In: *Advances in Neural Information Processing Systems*. 4520–4528.

Jin, C., Y. Zhang, S. Balakrishnan, M. J. Wainwright, and M. I. Jordan. (2016b). “Local maxima in the likelihood of Gaussian mixture models: Structural results and algorithmic consequences”. In: *Advances in Neural Information Processing Systems*. 4116–4124.

Jin, J. (2015). “Fast community detection by SCORE”. *The Annals of Statistics*. 43(1): 57–89.

Jin, J., Z. T. Ke, and W. Wang. (2017). “Phase transitions for high dimensional clustering and related problems”. *The Annals of Statistics*. 45(5): 2151–2189.

Johnstone, I. M. (2001). “On the distribution of the largest eigenvalue in principal components analysis”. *The Annals of statistics*. 29(2): 295–327.

Johnstone, I. M. and A. Y. Lu. (2009). “On consistency and sparsity for principal components analysis in high dimensions”. *Journal of the American Statistical Association*. 104(486): 682–693.

Johnstone, I. M. and D. Paul. (2018). “PCA in high dimensions: An orientation”. *Proceedings of the IEEE*. 106(8): 1277–1292.

Jolliffe, I. T. (1986). “Principal components in regression analysis”. In: *Principal component analysis*. Springer. 129–155.

Kalai, A. T., A. Moitra, and G. Valiant. (2010). “Efficiently learning mixtures of two Gaussians”. In: *Proceedings of the forty-second ACM symposium on Theory of computing*. 553–562.

Kannan, R., H. Salmasian, and S. Vempala. (2008). “The spectral method for general mixture models”. *SIAM Journal on Computing*. 38(3): 1141–1156.

Kannan, R. and S. Vempala. (2009). *Spectral algorithms*. Now Publishers Inc.

Karger, D. R., S. Oh, and D. Shah. (2013). “Efficient crowdsourcing for multi-class labeling”. In: *ACM SIGMETRICS/international conference on Measurement and modeling of computer systems*. 81–92.

Karger, D. R., S. Oh, and D. Shah. (2014). “Budget-optimal task allocation for reliable crowdsourcing systems”. *Operations Research*. 62(1): 1–24.

Kato, T. (2013). *Perturbation theory for linear operators*. Vol. 132. Springer Science & Business Media.

Ke, Z. T. and M. Wang. (2017). “A new SVD approach to optimal topic estimation”. *arXiv preprint arXiv:1704.07016*.

Keshavan, R., A. Montanari, and S. Oh. (2009). “Matrix completion from noisy entries”. In: *Advances in Neural Information Processing Systems*. 952–960.

Keshavan, R. H., A. Montanari, and S. Oh. (2010). “Matrix completion from a few entries”. *IEEE Transactions on Information Theory*. 56(6): 2980–2998.

Klopp, O. (2014). “Noisy low-rank matrix completion with general sampling distribution”. *Bernoulli*. 20(1): 282–303.

Kolda, T. G. and B. W. Bader. (2009). “Tensor decompositions and applications”. *SIAM review*. 51(3): 455–500.

Koltchinskii, V., M. Löffler, and R. Nickl. (2020). “Efficient estimation of linear functionals of principal components”. *The Annals of Statistics*. 48(1): 464–490.

Koltchinskii, V. and K. Lounici. (2016). “Asymptotics and concentration bounds for bilinear forms of spectral projectors of sample covariance”. In: *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*. Vol. 52. No. 4. 1976–2013.

Koltchinskii, V., K. Lounici, and A. B. Tsybakov. (2011). “Nuclear-norm penalization and optimal rates for noisy low-rank matrix completion”. *The Annals of Statistics*. 39(5): 2302–2329.

Koltchinskii, V. and D. Xia. (2016). “Perturbation of linear forms of singular vectors under Gaussian noise”. In: *High Dimensional Probability VII*. Springer. 397–423.

Kong, W., R. Somani, Z. Song, S. Kakade, and S. Oh. (2020). “Meta-learning for mixed linear regression”. In: *International Conference on Machine Learning*. 5394–5404.

Krahmer, F. and D. Stöger. (2021). “On the convex geometry of blind deconvolution and matrix completion”. *Communications on Pure and Applied Mathematics*. 74(4): 790–832.

Kreimer, N., A. Stanton, and M. D. Sacchi. (2013). “Tensor completion based on nuclear norm minimization for 5D seismic data reconstruction”. *Geophysics*. 78(6): V273–V284.

Kumar, A. and R. Kannan. (2010). “Clustering with spectral norm and the k -means algorithm”. In: *Proceedings of 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 299–308.

Kwon, J., N. Ho, and C. Caramanis. (2021). “On the minimax optimality of the EM algorithm for learning two-component mixed linear regression”. In: *International Conference on Artificial Intelligence and Statistics*. PMLR. 1405–1413.

Lawley, D. N. and A. E. Maxwell. (1962). “Factor analysis as a statistical method”. *Journal of the Royal Statistical Society. Series D (The Statistician)*. 12(3): 209–229.

Le, C. M. and E. Levina. (2015). “Estimating the number of communities in networks by spectral methods”. *arXiv preprint arXiv:1507.00827*.

Le, C. M., E. Levina, and R. Vershynin. (2017). “Concentration and regularization of random graphs”. *Random Structures & Algorithms*. 51(3): 538–561.

Le, C. M., E. Levina, and R. Vershynin. (2018). “Concentration of random graphs and application to community detection”. In: *Proceedings of the International Congress of Mathematicians: Rio de Janeiro 2018*. World Scientific. 2925–2943.

Lee, K., F. Krahmer, and J. Romberg. (2018). “Spectral methods for passive imaging: Nonasymptotic performance and robustness”. *SIAM Journal on Imaging Sciences*. 11(3): 2110–2164.

Lee, K., Y. Wu, and Y. Bresler. (2017). “Near-optimal compressed sensing of a class of sparse low-rank matrices via sparse power factorization”. *IEEE Transactions on Information Theory*. 64(3): 1666–1698.

Lei, J. and A. Rinaldo. (2015). “Consistency of spectral clustering in stochastic block models”. *The Annals of Statistics*. 43(1): 215–237.

Lei, J. and L. Zhu. (2014). “A generic sample splitting approach for refined community recovery in stochastic block models”. *arXiv preprint arXiv:1411.1469*.

Lei, L. (2019). “Unified $\ell_{2 \rightarrow \infty}$ eigenspace perturbation theory for symmetric random matrices”. *arXiv preprint arXiv:1909.04798*.

Lei, L., P. J. Bickel, and N. El Karoui. (2018). “Asymptotics for high dimensional regression M-estimates: Fixed design results”. *Probability Theory and Related Fields*. 172(3-4): 983–1079.

Li, B. (2018). *Sufficient dimension reduction: Methods and applications with R*. CRC Press.

Li, G., C. Cai, Y. Gu, H. V. Poor, and Y. Chen. (2021a). “Minimax estimation of linear functions of eigenvectors in the face of small eigen-gaps”. *arXiv preprint arXiv:2104.03298*.

Li, G. and Y. Gu. (2021). “Theory of spectral method for union of subspaces-based random geometry graph”. In: *International Conference on Machine Learning*. PMLR. 6337–6345.

Li, G., Y. Wei, Y. Chi, Y. Gu, and Y. Chen. (2020a). “Breaking the sample size barrier in model-based reinforcement learning with a generative model”. In: *Advances in Neural Information Processing Systems (NeurIPS)*. 12861–12872.

Li, J. and A. O. Hero. (2004). “A fast spectral method for active 3D shape reconstruction”. *Journal of Mathematical Imaging and Vision*. 20(1-2): 73–87.

Li, K.-C. (1992). “On principal Hessian directions for data visualization and dimension reduction: Another application of Stein’s lemma”. *Journal of the American Statistical Association*. 87(420): 1025–1039.

Li, R.-C. (1998). “Relative perturbation theory: II. Eigenspace and singular subspace variations”. *SIAM Journal on Matrix Analysis and Applications*. 20(2): 471–492.

Li, X. and V. Voroninski. (2013). “Sparse signal recovery from quadratic measurements via convex programming”. *SIAM Journal on Mathematical Analysis*. 45(5): 3019–3033.

Li, X., Y. Chen, and J. Xu. (2021b). “Convex relaxation methods for community detection”. *Statistical Science*. 36(1): 2–15.

Li, X., Y. Li, S. Ling, T. Strohmer, and K. Wei. (2020b). “When do birds of a feather flock together? k -means, proximity, and conic programming”. *Mathematical Programming*. 179(1-2): 295–341.

Li, X., S. Ling, T. Strohmer, and K. Wei. (2019). “Rapid, robust, and reliable blind deconvolution via nonconvex optimization”. *Applied and Computational Harmonic Analysis*. 47(3): 893–934.

Li, Y., K. Lee, and Y. Bresler. (2018). “Blind gain and phase calibration via sparse spectral methods”. *IEEE Transactions on Information Theory*. 65(5): 3097–3123.

Li, Y., Y. Chi, H. Zhang, and Y. Liang. (2020c). “Non-convex low-rank matrix recovery with arbitrary outliers via median-truncated gradient descent”. *Information and Inference: A Journal of the IMA*. 9(2): 289–325.

Li, Y., C. Ma, Y. Chen, and Y. Chi. (2021c). “Nonconvex matrix factorization from rank-one measurements”. *IEEE Transactions on Information Theory*. 67(3): 1928–1950.

Lin, N., R. Jing, Y. Wang, E. Yonekura, J. Fan, and L. Xue. (2017). “A statistical investigation of the dependence of tropical cyclone intensity change on the surrounding environment”. *Monthly Weather Review*. 145(7): 2813–2831.

Ling, S. (2020). “Near-optimal performance bounds for orthogonal and permutation group synchronization via spectral methods”. *arXiv preprint arXiv:2008.05341*.

Ling, S. and T. Strohmer. (2019). “Regularized gradient descent: A non-convex recipe for fast joint blind deconvolution and demixing”. *Information and Inference: A Journal of the IMA*. 8(1): 1–49.

Liu, A. and A. Moitra. (2020). “Tensor Completion Made Practical”. *Neural Information Processing Systems*.

Liu, J., P. Musialski, P. Wonka, and J. Ye. (2012). “Tensor completion for estimating missing values in visual data”. *IEEE transactions on Pattern Analysis and Machine Intelligence*. 35(1): 208–220.

Liu, L. T., E. Dobriban, and A. Singer. (2018). “ e PCA: High dimensional exponential family PCA”. *The Annals of Applied Statistics*. 12(4): 2121–2150.

Liu, Y. (2012). “Perturbation bounds for the stationary distributions of Markov chains”. *SIAM Journal on Matrix Analysis and Applications*. 33(4): 1057–1074.

Lloyd, S. (1982). “Least squares quantization in PCM”. *IEEE Transactions on Information Theory*. 28(2): 129–137.

Löffler, M., A. Y. Zhang, and H. H. Zhou. (2021). “Optimality of spectral clustering for Gaussian mixture model”. *accepted to The Annals of Statistics*.

Loh, P.-L. and M. J. Wainwright. (2012). “High-dimensional regression with noisy and missing data: provable guarantees with nonconvexity”. *The Annals of Statistics*. 40(3): 1637–1664.

Lounici, K. (2014). “High-dimensional covariance matrix estimation with missing observations”. *Bernoulli*. 20(3): 1029–1058.

Lu, Y. and H. H. Zhou. (2016). “Statistical and computational guarantees of Lloyd’s algorithm and its variants”. *arXiv preprint arXiv:1612.02099*.

Lu, Y. M. and G. Li. (2020). “Phase transitions of spectral initialization for high-dimensional non-convex estimation”. *Information and Inference: A Journal of the IMA*. 9(3): 507–541.

Luce, R. D. (2012). *Individual choice behavior: A theoretical analysis*. Courier Corporation.

Luo, W., W. Alghamdi, and Y. M. Lu. (2019). “Optimal spectral initialization for signal recovery with applications to phase retrieval”. *IEEE Transactions on Signal Processing*. 67(9): 2347–2356.

Lusseau, D., K. Schneider, O. J. Boisseau, P. Haase, E. Slooten, and S. M. Dawson. (2003). “The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations”. *Behavioral Ecology and Sociobiology*. 54(4): 396–405.

Ma, C., Y. Li, and Y. Chi. (2021a). “Beyond Procrustes: Balancing-free gradient descent for asymmetric low-rank matrix sensing”. *IEEE Transactions on Signal Processing*. 69: 867–877.

Ma, C., K. Wang, Y. Chi, and Y. Chen. (2020). “Implicit regularization in nonconvex statistical estimation: Gradient descent converges linearly for phase retrieval, matrix completion, and blind deconvolution”. *Foundations of Computational Mathematics*. 20(3): 451–632.

Ma, J., R. Dudeja, J. Xu, A. Maleki, and X. Wang. (2021b). “Spectral method for phase retrieval: an expectation propagation perspective”. *IEEE Transactions on Information Theory*. 67(2): 1332–1355.

Ma, J., J. Xu, and A. Maleki. (2019). “Optimization-based AMP for Phase Retrieval: The Impact of Initialization and ℓ_2 -regularization”. *IEEE Transactions on Information Theory*. 65(6): 3600–3629.

Ma, Z. (2013). “Sparse principal component analysis and iterative thresholding”. *The Annals of Statistics*. 41(2): 772–801.

MacQueen, J. (1967). “Some methods for classification and analysis of multivariate observations”. In: *Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability*. Vol. 1. No. 14. Oakland, CA, USA. 281–297.

Mahoney, M. W. (2016). “Lecture notes on randomized linear algebra”. *arXiv preprint arXiv:1608.04481*.

Massoulié, L. (2014). “Community detection thresholds and the weak Ramanujan property”. In: *Symposium on Theory of computing*. 694–703.

Maunu, T., T. Zhang, and G. Lerman. (2019). “A well-tempered landscape for non-convex robust subspace recovery”. *Journal of Machine Learning Research*. 20(37): 1–59.

McCrae, R. R. and O. P. John. (1992). “An introduction to the five-factor model and its applications”. *Journal of personality*. 60(2): 175–215.

McRae, A. D. and M. A. Davenport. (2021). “Low-rank matrix completion and denoising under Poisson noise”. *Information and Inference: A Journal of the IMA*. 10(2): 697–720.

McSherry, F. (2001). “Spectral partitioning of random graphs”. In: *IEEE Symposium on Foundations of Computer Science*. 529–537.

Mezard, M. and A. Montanari. (2009). *Information, Physics, and Computation*. Oxford University Press.

Michoel, T. and B. Nachtergael. (2012). “Alignment and integration of complex networks by hypergraph-based spectral clustering”. *Physical Review E*. 86(5): 056111.

Mitrophanov, A. Y. (2005). “Sensitivity and convergence of uniformly ergodic Markov chains”. *Journal of Applied Probability*. 42(4): 1003–1014.

Mixon, D. G., S. Villar, and R. Ward. (2017). “Clustering subgaussian mixtures by semidefinite programming”. *Information and Inference: A Journal of the IMA*. 6(4): 389–415.

Moitra, A. and A. S. Wein. (2019). “Spectral methods from tensor networks”. In: *ACM SIGACT Symposium on Theory of Computing*. 926–937.

Monardo, V. and Y. Chi. (2019). “On the sensitivity of spectral initialization for noisy phase retrieval”. In: *International Conference on Acoustics, Speech and Signal Processing*. 5172–5176.

Mondelli, M. and A. Montanari. (2019). “Fundamental limits of weak recovery with applications to phase retrieval”. *Foundations of Computational Mathematics*. 19(3): 703–773.

Montanari, A. (2011). “EE 378B: Statistical Signal Processing”. *Lecture notes, Stanford University*.

Montanari, A. and N. Sun. (2018). “Spectral algorithms for tensor completion”. *Communications on Pure and Applied Mathematics*. 71(11): 2381–2425.

Mossel, E., J. Neeman, and A. Sly. (2015). “Consistency thresholds for the planted bisection model”. In: *Symposium on Theory of computing*. 69–75.

Nadler, B. (2008). “Finite sample approximation results for principal component analysis: A matrix perturbation approach”. *The Annals of Statistics*. 36(6): 2791–2817.

Nayer, S., P. Narayanamurthy, and N. Vaswani. (2019). “Phaseless PCA: Low-rank matrix recovery from column-wise phaseless measurements”. In: *International Conference on Machine Learning*. 4762–4770.

Ndaoud, M. (2018). “Sharp optimal recovery in the two component Gaussian mixture model”. *arXiv preprint arXiv:1812.08078*.

Negahban, S. and M. J. Wainwright. (2011). “Estimation of (near) low-rank matrices with noise and high-dimensional scaling”. *The Annals of Statistics*. 39(2): 1069–1097.

Negahban, S., S. Oh, and D. Shah. (2016). “Rank centrality: Ranking from pairwise comparisons”. *Operations Research*. 65(1): 266–287.

Negahban, S. and M. J. Wainwright. (2012). “Restricted strong convexity and weighted matrix completion: Optimal bounds with noise”. *Journal of Machine Learning Research*. 13(1): 1665–1697.

Netrapalli, P., P. Jain, and S. Sanghavi. (2015). “Phase retrieval using alternating minimization”. *IEEE Transactions on Signal Processing*. 18(63): 4814–4826.

Netrapalli, P., U. Niranjan, S. Sanghavi, A. Anandkumar, and P. Jain. (2014). “Non-convex robust PCA”. In: *Advances in Neural Information Processing Systems*. 1107–1115.

Newman, M. E. (2006). “Finding community structure in networks using the eigenvectors of matrices”. *Physical review E*. 74(3): 036104.

Newman, M. E. (2013). “Spectral methods for community detection and graph partitioning”. *Physical Review E*. 88(4): 042822.

Ng, A. Y., M. I. Jordan, and Y. Weiss. (2002). “On spectral clustering: Analysis and an algorithm”. In: *Advances in Neural Information Processing Systems*. 849–856.

Oh, S. and D. Shah. (2014). “Learning mixed multinomial logit model from ordinal data”. In: *Advances in Neural Information Processing Systems*. 595–603.

Oliveira, R. I. (2009). “Concentration of the adjacency matrix and of the Laplacian in random graphs with independent edges”. *arXiv preprint arXiv:0911.0600*.

Olver, P. J., C. Shakiban, and C. Shakiban. (2006). *Applied linear algebra*. Vol. 1. Springer.

Oymak, S., A. Jalali, M. Fazel, Y. C. Eldar, and B. Hassibi. (2015). “Simultaneously structured models with application to sparse and low-rank matrices”. *IEEE Transactions on Information Theory*. 61(5): 2886–2908.

Pachauri, D., R. Kondor, and V. Singh. (2013). “Solving the multi-way matching problem by permutation synchronization”. *Advances in Neural Information Processing Systems*. 26: 1860–1868.

Page, L., S. Brin, R. Motwani, and T. Winograd. (1999). “The PageRank citation ranking: Bringing order to the web.” *Tech. rep.* Stanford InfoLab.

Pananjady, A. and M. J. Wainwright. (2020). “Instance-dependent ℓ_∞ -bounds for policy evaluation in tabular reinforcement learning”. *IEEE Transactions on Information Theory*. 67(1): 566–585.

Paul, D. (2007). “Asymptotics of sample eigenstructure for a large dimensional spiked covariance model”. *Statistica Sinica*: 1617–1642.

Pearson, K. (1894). “Contributions to the mathematical theory of evolution”. *Philosophical Transactions of the Royal Society of London*. A. 185: 71–110.

Peng, J. and Y. Wei. (2007). “Approximating k -means-type clustering via semidefinite programming”. *SIAM Journal on Optimization*. 18(1): 186–205.

Perry, A., A. S. Wein, A. S. Bandeira, and A. Moitra. (2018). “Optimality and sub-optimality of PCA I: Spiked random matrix models”. *The Annals of Statistics*. 46(5): 2416–2451.

Qu, Q., Y. Zhang, Y. C. Eldar, and J. Wright. (2019). “Convolutional phase retrieval via gradient descent”. *IEEE Transactions on Information Theory*. 66(3): 1785–1821.

Raginsky, M. and I. Sason. (2013). “Concentration of measure inequalities in information theory, communications, and coding”. *Foundations and Trends in Communications and Information Theory*. 10(1-2): 1–250.

Rajkumar, A. and S. Agarwal. (2014). “A statistical convergence perspective of algorithms for rank aggregation from pairwise data”. In: *International Conference on Machine Learning*. 118–126.

Reynolds, D. A. and R. C. Rose. (1995). “Robust text-independent speaker identification using Gaussian mixture speaker models”. *IEEE Transactions on Speech and Audio Processing*. 3(1): 72–83.

Richard, E. and A. Montanari. (2014). “A statistical model for tensor PCA”. In: *Advances in Neural Information Processing Systems*. 2897–2905.

Rohe, K., S. Chatterjee, and B. Yu. (2011). “Spectral clustering and the high-dimensional stochastic blockmodel”. *The Annals of Statistics*. 39(4): 1878–1915.

Rudolf, D. and N. Schweizer. (2018). “Perturbation theory for Markov chains via Wasserstein distance”. *Bernoulli*. 24(4A): 2610–2639.

Salehi, F., E. Abbasi, and B. Hassibi. (2018a). “A precise analysis of phasemax in phase retrieval”. In: *2018 IEEE International Symposium on Information Theory (ISIT)*. IEEE. 976–980.

Salehi, F., E. Abbasi, and B. Hassibi. (2018b). “Learning without the phase: Regularized phasemax achieves optimal sample complexity”. In: *Advances in Neural Information Processing Systems*. 8641–8652.

Sanghavi, S., R. Ward, and C. D. White. (2017). “The local convexity of solving systems of quadratic equations”. *Results in Mathematics*. 71(3-4): 569–608.

Sarkar, P. and P. J. Bickel. (2015). “Role of normalization in spectral clustering for stochastic blockmodels”. *The Annals of Statistics*. 43(3): 962–990.

Scharf, L. L. (1991). “The SVD and reduced rank signal processing”. *Signal Processing*. 25(2): 113–133.

Schönemann, P. H. (1966). “A generalized solution of the orthogonal procrustes problem”. *Psychometrika*. 31(1): 1–10.

Seeley, J. R. (1949). “The net of reciprocal influence. a problem in treating sociometric data”. *Canadian Journal of Experimental Psychology*. 3: 234.

Shah, N., S. Balakrishnan, A. Guntuboyina, and M. Wainwright. (2016). “Stochastically transitive models for pairwise comparisons: Statistical and computational issues”. In: *International Conference on Machine Learning*. 11–20.

Shah, N. B., S. Balakrishnan, and M. J. Wainwright. (2019). “Feeling the Bern: Adaptive estimators for Bernoulli probabilities of pairwise comparisons”. *IEEE Transactions on Information Theory*. 65(8): 4854–4874.

Shah, N. B., J. K. Bradley, A. Parekh, M. Wainwright, and K. Ramchandran. (2013). “A case for ordinal peer-evaluation in MOOCs”. In: *NIPS Workshop on Data Driven Education*. 1–8.

Shapiro, A., Y. Xie, and R. Zhang. (2018). “Matrix completion with deterministic pattern: A geometric perspective”. *IEEE Transactions on Signal Processing*. 67(4): 1088–1103.

Shechtman, Y., A. Beck, and Y. C. Eldar. (2014). “GESPAR: Efficient phase retrieval of sparse signals”. *IEEE Transactions on Signal Processing*. 62(4): 928–938.

Shechtman, Y., Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev. (2015). “Phase retrieval with application to optical imaging: a contemporary overview”. *IEEE Signal Processing Magazine*. 32(3): 87–109.

Shen, Y., Q. Huang, N. Srebro, and S. Sanghavi. (2016). “Normalized spectral map synchronization”. In: *Advances in Neural Information Processing Systems*. 4925–4933.

Shi, J. and J. Malik. (2000). “Normalized cuts and image segmentation”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 22(8): 888–905.

Sidiropoulos, N. D., L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos. (2017). “Tensor decomposition for signal processing and machine learning”. *IEEE Transactions on Signal Processing*. 65(13): 3551–3582.

Singer, A. (2011). “Angular synchronization by eigenvectors and semidefinite programming”. *Applied and Computational Harmonic Analysis*. 30(1): 20–36.

Singer, A. and Y. Shkolnisky. (2011). “Three-dimensional structure determination from common lines in cryo-EM by eigenvectors and semidefinite programming”. *SIAM Journal on Imaging Sciences*. 4(2): 543–572.

Soltanolkotabi, M. (2019). “Structured signal recovery from quadratic measurements: Breaking sample complexity barriers via nonconvex optimization”. *IEEE Transactions on Information Theory*. 65(4): 2374–2400.

Srivastava, P. R., P. Sarkar, and G. A. Hanasusanto. (2019). “A robust spectral clustering algorithm for sub-Gaussian mixture models with outliers”. *arXiv preprint arXiv:1912.07546*.

Stein, C. (1972). “A bound for the error in the normal approximation to the distribution of a sum of dependent random variables”. In: *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*.

Stewart, G. W. and J.-G. Sun. (1990). *Matrix perturbation theory*. Academic Press.

Stock, J. H. and M. W. Watson. (2002). “Forecasting using principal components from a large number of predictors”. *Journal of the American statistical association*. 97(460): 1167–1179.

Stock, J. H. and M. W. Watson. (2016). “Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics”. In: *Handbook of macroeconomics*. Vol. 2. Elsevier. 415–525.

Sun, J. (1987). *Matrix perturbation analysis (in Chinese)*. Vol. 6. Science Press.

Sun, R. and Z.-Q. Luo. (2016). “Guaranteed matrix completion via non-convex factorization”. *IEEE Transactions on Information Theory*. 62(11): 6535–6579.

Sun, Y., Z. Liang, X. Huang, and Q. Huang. (2018). “Joint map and symmetry synchronization”. In: *European Conference on Computer Vision (ECCV)*. 251–264.

Sun, Y., J. Zhuo, A. Mohan, and Q. Huang. (2019). “ K -best transformation synchronization”. In: *IEEE International Conference on Computer Vision*. 10252–10261.

Sur, P. and E. J. Candès. (2019). “A modern maximum-likelihood theory for high-dimensional logistic regression”. *Proceedings of the National Academy of Sciences*. 116(29): 14516–14525.

Sur, P., Y. Chen, and E. J. Candès. (2019). “The likelihood ratio test in high-dimensional logistic regression is asymptotically a rescaled chi-square”. *Probability Theory and Related Fields*. 175(1-2): 487–558.

Tan, Y. S. and R. Vershynin. (2019). “Phase retrieval via randomized Kaczmarz: Theoretical guarantees”. *Information and Inference: A Journal of the IMA*. 8(1): 97–123.

Tao, T. (2012). *Topics in Random Matrix Theory. Graduate Studies in Mathematics*. Providence, Rhode Island: American Mathematical Society.

Titterington, D. M., A. F. Smith, and U. E. Makov. (1985). *Statistical analysis of finite mixture distributions*. Wiley.

Tomasi, C. and T. Kanade. (1992). “Shape and motion from image streams under orthography: a factorization method”. *International Journal of Computer Vision*. 9(2): 137–154.

Tong, T., C. Ma, and Y. Chi. (2021a). “Accelerating ill-conditioned low-rank matrix estimation via scaled gradient descent”. *Journal of Machine Learning Research*. 22(150): 1–63.

Tong, T., C. Ma, and Y. Chi. (2021b). “Low-rank matrix recovery with scaled subgradient methods: Fast and robust convergence without the condition number”. *IEEE Transactions on Signal Processing*. 69: 2396–2409.

Tong, T., C. Ma, A. Prater-Bennette, E. Tripp, and Y. Chi. (2021c). “Scaling and scalability: Provable nonconvex low-rank tensor estimation from incomplete measurements”. *arXiv preprint arXiv:2104.14526*.

Tripuraneni, N., C. Jin, and M. Jordan. (2021). “Provable meta-learning of linear representations”. In: *International Conference on Machine Learning*. PMLR. 10434–10443.

Tropp, J. (2011). “Freedman’s inequality for matrix martingales”. *Electronic Communications in Probability*. 16: 262–270.

Tropp, J. A. (2012). “User-friendly tail bounds for sums of random matrices”. *Foundations of Computational Mathematics*. 12(4): 389–434.

Tropp, J. A. (2015). “An introduction to matrix concentration inequalities”. *Foundations and Trends® in Machine Learning*. 8(1-2): 1–230.

Tsybakov, A. B. (2009). *Introduction to Nonparametric Estimation. Springer Series in Statistics*. Springer.

Tu, S., R. Boczar, M. Simchowitz, M. Soltanolkotabi, and B. Recht. (2016). “Low-rank solutions of linear matrix equations via procrustes flow”. In: *International Conference on Machine Learning*. 964–973.

Turk, M. A. and A. P. Pentland. (1991). “Face recognition using eigenfaces”. *Conference on Computer Vision and Pattern Recognition*: 586–591.

Van der Vaart, A. W. (2000). *Asymptotic statistics*. Vol. 3. Cambridge university press.

Vaswani, N. (2020). “Nonconvex structured phase retrieval: A focus on provably correct approaches”. *IEEE Signal Processing Magazine*. 37(5): 67–77.

Vaswani, N., Y. Chi, and T. Bouwmans. (2018). “Rethinking PCA for modern data sets: Theory, algorithms, and applications”. *Proceedings of the IEEE*. 106(8): 1274–1276.

Vaswani, N., S. Nayer, and Y. C. Eldar. (2017). “Low-rank phase retrieval”. *IEEE Transactions on Signal Processing*. 65(15): 4059–4074.

Vempala, S. and G. Wang. (2004). “A spectral algorithm for learning mixture models”. *Journal of Computer and System Sciences*. 68(4): 841–860.

Vershynin, R. (2017). *High-dimensional probability*. Cambridge University Press.

Vidal, R., Y. Ma, and S. S. Sastry. (2016). *Generalized principal component analysis*. Springer.

Vigna, S. (2016). “Spectral ranking”. *Network Science*. 4(4): 433–445.

Von Luxburg, U. (2007). “A tutorial on spectral clustering”. *Statistics and Computing*. 17(4): 395–416.

Vu, V. and J. Lei. (2012). “Minimax rates of estimation for sparse PCA in high dimensions”. In: *Artificial intelligence and statistics*. 1278–1286.

Vu, V. Q. and J. Lei. (2013). “Minimax sparse principal subspace estimation in high dimensions”. *The Annals of Statistics*. 41(6): 2905–2947.

Wainwright, M. J. (2019). *High-dimensional statistics: A non-asymptotic viewpoint*. Vol. 48. Cambridge University Press.

Wang, B., Y. Yan, and J. Fan. (2021a). “Sample-efficient reinforcement learning for linearly-parameterized MDPs with a generative model”. *arXiv preprint arXiv:2105.14016*.

Wang, G., G. B. Giannakis, and Y. C. Eldar. (2018a). “Solving systems of random quadratic equations via truncated amplitude flow”. *IEEE Transactions on Information Theory*. 64(2): 773–794.

Wang, G., L. Zhang, G. B. Giannakis, M. Akçakaya, and J. Chen. (2018b). “Sparse phase retrieval via truncated amplitude flow”. *IEEE Transactions on Signal Processing*. 66(2): 479–491.

Wang, H., J. Chen, and K. Wei. (2021b). “Entrywise convergence of Riemannian gradient method for low rank tensor completion via Tucker decomposition”. *arXiv preprint arXiv:2108.07899*.

Wang, W. and J. Fan. (2017). “Asymptotics of empirical eigenstructure for high dimensional spiked covariance”. *The Annals of Statistics*. 45(3): 1342.

Wang, Z., Y. Liang, and P. Ji. (2020). “Spectral algorithms for community detection in directed networks”. *Journal of Machine Learning Research*. 21(153): 1–45.

Wedin, P.-Å. (1972). “Perturbation bounds in connection with singular value decomposition”. *BIT Numerical Mathematics*. 12(1): 99–111.

Wei, K., J.-F. Cai, T. F. Chan, and S. Leung. (2016). “Guarantees of Riemannian optimization for low rank matrix recovery”. *SIAM Journal on Matrix Analysis and Applications*. 37(3): 1198–1222.

Wright, J. and Y. Ma. (2020). *High-Dimensional Data Analysis with Low-Dimensional Models: Principles, Computation, and Applications*. Cambridge University Press.

Wu, Y. and P. Yang. (2020). “Optimal estimation of Gaussian mixtures via denoised method of moments”. *The Annals of Statistics*. 48(4): 1981–2007.

Xia, D. (2021). “Normal approximation and confidence region of singular subspaces”. *Electronic Journal of Statistics*. 15(2): 3798–3851.

Xia, D. and M. Yuan. (2019). “On polynomial time methods for exact low-rank tensor completion”. *Foundations of Computational Mathematics*. 19(6): 1265–1313.

Xia, D. and M. Yuan. (2021). “Statistical inferences of linear forms for noisy matrix completion”. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 83(1): 58–77.

Xia, D., M. Yuan, and C.-H. Zhang. (2021). “Statistically optimal and computationally efficient low rank tensor completion from noisy entries”. *The Annals of Statistics*. 49(1): 76–99.

Xia, D., A. R. Zhang, and Y. Zhou. (2020). “Inference for low-rank tensors—No need to debias”. *arXiv preprint arXiv:2012.14844*.

Xia, Y., H. Tong, W. K. Li, and L.-X. Zhu. (2002). “An adaptive estimation of dimension reduction space”. *Journal of Royal Statistical Society Series B*. 64: 363–410.

Xie, L., Y. Xie, and G. V. Moustakides. (2020). “Sequential subspace change point detection”. *Sequential Analysis*. 39(3): 307–335.

Xu, J., D. J. Hsu, and A. Maleki. (2016). “Global analysis of expectation maximization for mixtures of two Gaussians”. In: *Advances in Neural Information Processing Systems*. 2676–2684.

Xu, J., A. Maleki, and K. R. Rad. (2019). “Consistent risk estimation in high-dimensional linear regression”. *arXiv preprint arXiv:1902.01753*.

Xu, L. and M. I. Jordan. (1996). “On convergence properties of the EM algorithm for Gaussian mixtures”. *Neural Computation*. 8(1): 129–151.

Yan, Y., Y. Chen, and J. Fan. (2021). “Inference for heteroskedastic PCA with missing data”. *arXiv preprint arXiv:2107.12365*.

Yang, Z., Z. Wang, H. Liu, Y. Eldar, and T. Zhang. (2016). “Sparse nonlinear regression: Parameter estimation under nonconvexity”. In: *International Conference on Machine Learning*. 2472–2481.

Yang, Z., L. F. Yang, E. X. Fang, T. Zhao, Z. Wang, and M. Neykov. (2019). “Misspecified nonconvex statistical optimization for sparse phase retrieval”. *Mathematical Programming*. 176(1-2): 545–571.

Yi, X., C. Caramanis, and S. Sanghavi. (2014). “Alternating minimization for mixed linear regression”. In: *International Conference on Machine Learning*. 613–621.

Yi, X., D. Park, Y. Chen, and C. Caramanis. (2016). “Fast algorithms for robust PCA via gradient descent”. In: *Advances in Neural Information Processing Systems*. 4152–4160.

Yu, Y., T. Wang, and R. J. Samworth. (2015). “A useful variant of the Davis-Kahan theorem for statisticians”. *Biometrika*. 102(2): 315–323.

Yuan, Y. and A. Qu. (2018). “Community detection with dependent connectivity”. *arXiv preprint arXiv:1812.06406*.

Yuan, Z., H. Wang, and Q. Wang. (2019). “Phase retrieval via sparse Wirtinger flow”. *Journal of Computational and Applied Mathematics*. 355: 162–173.

Zhang, A. Y. and H. H. Zhou. (2020). “Theoretical and computational guarantees of mean field variational inference for community detection”. *The Annals of Statistics*. 48(5): 2575–2598.

Zhang, A., T. T. Cai, and Y. Wu. (2021). “Heteroskedastic PCA: Algorithm, optimality, and applications”. *accepted to The Annals of Statistics*.

Zhang, A. and M. Wang. (2020). “Spectral state compression of Markov processes”. *IEEE Transactions on Information Theory*. 66(5): 3202–3231.

Zhang, A. and D. Xia. (2018). “Tensor SVD: Statistical and computational limits”. *IEEE Transactions on Information Theory*. 64(11): 7311–7338.

Zhang, A. R., Y. Luo, G. Raskutti, and M. Yuan. (2020a). “ISLET: Fast and optimal low-rank tensor regression via importance sketching”. *SIAM Journal on Mathematics of Data Science*. 2(2): 444–479.

Zhang, H., Y. Chi, and Y. Liang. (2016). “Provable non-convex phase retrieval with outliers: Median truncated Wirtinger flow”. In: *International Conference on Machine Learning*. 1022–1031.

Zhang, H., Y. Chi, and Y. Liang. (2018a). “Median-truncated nonconvex approach for phase retrieval with outliers”. *IEEE Transactions on Information Theory*. 64(11): 7287–7310.

Zhang, H., Y. Zhou, Y. Liang, and Y. Chi. (2017). “A nonconvex approach for phase retrieval: Reshaped Wirtinger flow and incremental algorithms”. *The Journal of Machine Learning Research*. 18(1): 5164–5198.

Zhang, K., S. Kakade, T. Basar, and L. Yang. (2020b). “Model-based multi-agent RL in zero-sum Markov games with near-optimal sample complexity”. *Advances in Neural Information Processing Systems*. 33.

Zhang, L., G. Wang, G. B. Giannakis, and J. Chen. (2018b). “Compressive phase retrieval via reweighted amplitude flow”. *IEEE Transactions on Signal Processing*. 66(19): 5029–5040.

Zhang, X., S. Du, and Q. Gu. (2018c). “Fast and sample efficient inductive matrix completion via multi-phase Procrustes flow”. In: *International Conference on Machine Learning*. 5756–5765.

Zhang, X., L. Wang, Y. Yu, and Q. Gu. (2018d). “A primal-dual analysis of global optimality in nonconvex low-rank matrix recovery”. In: *International Conference on Machine Learning*. 5857–5866.

Zhang, Y., E. Levina, and J. Zhu. (2020c). “Detecting overlapping communities in networks using spectral methods”. *SIAM Journal on Mathematics of Data Science*. 2(2): 265–283.

Zhang, Y., X. Chen, D. Zhou, and M. I. Jordan. (2014). “Spectral methods meet EM: A provably optimal algorithm for crowdsourcing”. In: *Advances in Neural Information Processing Systems*. 1260–1268.

Zhang, Y., Q. Qu, and J. Wright. (2020d). “From symmetry to geometry: Tractable nonconvex problems”. *arXiv preprint arXiv:2007.06753*.

Zhao, Y., E. Levina, and J. Zhu. (2012). “Consistency of community detection in networks under degree-corrected stochastic block models”. *The Annals of Statistics*. 40(4): 2266–2292.

Zheng, Q. and J. Lafferty. (2015). “A convergent gradient descent algorithm for rank minimization and semidefinite programming from random linear measurements”. In: *Advances in Neural Information Processing Systems*. 109–117.

Zheng, Q. and J. Lafferty. (2016). “Convergence analysis for rectangular matrix completion using Burer-Monteiro factorization and gradient descent”. *arXiv preprint arXiv:1605.07051*.

Zhong, K., Z. Song, P. Jain, P. L. Bartlett, and I. S. Dhillon. (2017). “Recovery guarantees for one-hidden-layer neural networks”. In: *International Conference on Machine Learning*. 4140–4149.

Zhong, Y. and N. Boumal. (2018). “Near-optimal bounds for phase synchronization”. *SIAM Journal on Optimization*. 28(2): 989–1016.

Zhu, Z., Y. Wang, D. Robinson, D. Naiman, R. Vidal, and M. Tsakiris. (2018). “Dual principal component pursuit: Improved analysis and efficient algorithms”. In: *Advances in Neural Information Processing Systems*. 2171–2181.

Zhu, Z., T. Wang, and R. J. Samworth. (2019). “High-dimensional principal component analysis with heterogeneous missingness”. *arXiv preprint arXiv:1906.12125*.

Zou, H., T. Hastie, and R. Tibshirani. (2006). “Sparse principal component analysis”. *Journal of Computational and Graphical Statistics*. 15(2): 265–286.