TENSOR DATA ANALYSIS: STATISTICAL METHODS WITH LOW-DIMENSIONAL STRUCTURES AND REAL-WORLD APPLICATIONS

by

Jiaxin Hu

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Statistics)

at the

UNIVERSITY OF WISCONSIN–MADISON

2025

Date of final oral examination: 04/17/2025

The dissertation is approved by the following members of the Final Oral Committee:

Miaoyan Wang, Associate Professor, Statistics
Karl Rohe, Professor, Statistics
Daniel I. Bolnick, Professor, Ecology and Evolutionary Biology
Joshua R. Cape, Assistant Professor, Statistics
Sündüz Keleş, Professor, Statistics & Biostatistics and Medical Informatics

ACKNOWLEDGMENTS

During my Ph.D. journey, I have received countless support and encouragement from many people around me. Without them, I could not have completed my Ph.D. and achieved several meaningful milestones along the way. I would like to take this opportunity to express my deep gratitude and sincere appreciation.

I would like to express my deepest gratitude to my advisor, Prof. Miaoyan Wang, for her exceptional mentorship and unwavering support throughout my graduate studies. I vividly remember February 14, 2019 — our first research meeting in Room 1250 of the Medical Sciences Center — when I first encountered the word “tensor". She introduced tensor analysis through a movie rating example, where the data forms a three-dimensional array representing ratings by different users, for different movies, at different times. To this day, I still use that example when explaining tensors to others. She opened the door for me to this fascinating world of tensors and shaped my academic focus in statistical methodology. Under her guidance, I received comprehensive training — from developing a high-level academic perspective to handling the intricate details of rigorous research. She taught me how to extract intuition and build understanding from complex ideas, as well as how to overcome research bottlenecks and write papers in a professional tone. She offered all kinds of support, both within and beyond research projects. The knowledge and skills I gained from her will benefit me for a lifetime. I could not be more appreciative or more fortunate to have had her as my advisor.

I would like to extend my heartfelt thanks to my dissertation committee members. I sincerely thank Prof. Daniel I. Bolnick for his strong support, patient guidance, and outstanding biological expertise coupled with statistical insight during our collaboration. I thank Prof. Joshua R. Cape for his thoughtful discussions and detailed comments on my research paper. I thank Prof. Karl Rohe and Prof. Sündüz Keleş for their valuable and inspiring feedback on my defense work. I am genuinely grateful to my committee for attending my defense and helping me improve my dissertation.

I would like to acknowledge my collaborators and all those who have supported

my research work. I thank Prof. Jesse Weber for his genetics expertise and biological feedback to the stickleback project. I thank Prof. Hyunseung Kang and Prof. Nicolás García Trillos for generously answering my random research questions at the office and Aldo's Café. I thank my Ph.D. cohorts Xinran Miao, Runshi Tang, and Kwangmoon Park for their help in the preparation of my solo paper. I give special thanks to Chanwoo Lee, my Ph.D. collaborator and academic brother, for his dedication to my tensor projects and his insightful suggestions on both academic and career matters.

In addition, I would also like to thank all my friends who have given me enormous encouragement and support. I have made many good friends at Madison, and the number of name is too large to show the full list. Particularly, I would like to thank the members of our fitness group, Yuchen Zeng, Ting Cai, and Xinran Miao, who encouraged me to step outside my apartment during my most stressful period. I would like to thank my cohorts Kwangmoon Park, Zhongling Jonquil Liao, Chenghui Zheng, Zexuan Sun, Jingqi Duan, Siqi Shen, Jitian Zhao, Chenhui Li, Sixu Li, and Chaoran Wang for the enjoyable and relaxing chats about research and life. I would also like to thank Chanwoo Lee and Siyu Wang for the relieving and informative conversations as senior friends. Meanwhile, special thanks go to my close friends far from Madison. Special thanks to Ran Tu, my best friend since the first day of undergrad, who has always been there with great comfort and encouragement. And special thanks to Haiming Ning for the immense companionship and care that supported me through difficult times. I sincerely appreciate all my friends, even those not mentioned here. The moments shared with friends have made my Madison journey more colorful and memorable.

Last but not least, I would like to particularly thank my parents, Zhensheng Hu and Chunhui Hu, for their support. It is their endless, non-stopping and unconditional support - love, care, trust, help, and encouragement - that made me who I am today.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The prevalence of multiway array data across diverse domains such as social networks, neuroscience, and genomics has brought tensor data analysis to the forefront of statistical research. Tensors effectively capture complex multi-dimensional structures beyond the scope of matrix-based methods. Recently, a growing body of statistical methods leveraging low-dimensional tensor structures has been proposed for tasks including decomposition, clustering, and hypothesis testing. However, tensor methods still remain less developed than their matrix counterparts. Methodologically, the inherent multilinear structure of tensors introduces unique theoretical challenges, while in practice, issues related to model specification and computational scalability often arise. In this thesis, we introduce four statistical methods based on low-dimensional tensor structures, aiming to advance both methodology and practical application in tensor data analysis.

In Chapter 2, we develop a *supervised tensor decomposition* (SupTD) method that incorporates multiple feature matrices as side information. Unlike most unsupervised decomposition methods designed for a single tensor observation, SupTD captures the effective dimension reduction of the data tensor confined to feature space of interest. SupTD also handles a broad range of data types, including continuous, count, and binary observations. Applications in neuroimaging and social science data demonstrate the efficacy of SupTD.

Chapters 3 and 4 both tackle the multiway clustering problem, but have different focuses. In Chapter 3, we consider the clustering under the *degree-corrected tensor block model* (dTBM), allowing unknown node degree heterogeneity. Theoretical phase transition of clustering performance and the statistical-computational behaviors are well studied under this dTBM setup. In Chapter 4, we pay a special attention to computational performance and propose a *large-scale multiway clustering framework under tensor block model* (LS-TBM). LS-TBM reduces the time and space complexities from polynomial to logarithmic rates while maintaining the exact recovery of community structures. Large-scale experiments with real large-scale Uber Pickup data highlight LS-TBM's superior performance in practice.

Finally, Chapter 5 presents a novel tensor-based method for hypothesis testing under a genetics context. We develop a *spectral network quantitative trait loci analysis* (snQTL), to map genetic loci affecting the gene co-expression network. Application of snQTL to three-spined sticklebacks reveals candidate loci missed by standard methods and estimates network changes led by genetic variants. snQTL provides a novel solution for comparing biological networks, opening doors for deeper network analysis in genetics.

1 INTRODUCTION

1.1 Motivation

Multiway arrays have been widely collected in various fields, including social networks (Anandkumar et al., 2014), neuroscience (Wang et al., 2017), and computer science (Koniusz and Cherian, 2016). Tensors effectively represent multiway data and serve as the foundation in higher-order data analysis. One data example is from a multi-tissue, multi-individual gene expression study (Wang et al., 2019; Hore et al., 2016), where the data tensor consists of expression measurements indexed by (gene, individual, tissue) triplets. Another example is a hypergraph network (Ghoshdastidar and Dukkipati, 2017b,a; Ahn et al., 2019; Ke et al., 2019) in social science. A K-uniform hypergraph can be naturally represented as an order-K tensor, where each entry indicates the presence of a K-way hyperedge among nodes (a.k.a. entities). As multiway arrays become increasingly available due to advances in data collection and storage techniques, tensor data analysis has come under the spotlight for its ability to uncover meaningful insights from high-dimensional and higher-order information across diverse applications.

A large branch of statistical methods for tensor data analysis (Kolda and Bader, 2009; Goldfarb and Qin, 2014; Auddy et al., 2024) utilizes low-dimensional structures to capture the critical patterns efficiently. Popular low-dimensional structures include low-rankness (Wang and Li, 2020a; Luo and Zhang, 2021, 2022; Hu et al., 2022) under different tensor decomposition models, and the checkerboard block structure (Wang and Zeng, 2019; Han et al., 2022a; Hu and Wang, 2023). The choice of structural assumptions often depends on the specific problem context and analytical goals. For instance, CANDECOMP/PARAFAC (CP) decomposition (Hitchcock, 1927) has been applied to tensor completion (Wang and Li, 2020a) and parameter estimation (Sun et al., 2017); Tucker decomposition (Tucker, 1966) has been employed for dimension reduction (Zhang and Xia, 2018) and regression (Hu et al., 2022); and checkerboard structures have primarily been used for clustering tasks (Wang and Zeng, 2019; Han et al., 2022a; Hu and Wang, 2023; Hu,

2025). Though some nonparametric methods (Lee and Wang, 2021, 2024) have been proposed for analyzing high-rank tensors, we focus on this popular class of methods based on low-dimensional assumptions due to their superior efficiency and interpretability.

Real-world applications for tensor data analysis have also grown in parallel with methodological advancements. On one hand, novel tensor techniques provide new analytical tools and perspectives for re-examining existing multiway datasets that were previously analyzed using matrix-based or lower-order approaches. For example, tensor-based methods (Hore et al., 2016; Wang et al., 2019) have revealed new genetic insights from multi-tissue gene expression data that beyond traditional association studies. On the other hand, innovations in scientific technologies have given rise to novel tensor-like datasets, further motivating the development of specialized tensor methods. For instance, a joint tensor analysis approach (Park and Keleş, 2024) has been proposed to handle single-cell 3D genome Hi-C data, enabling more effective modeling of the complex genome spatial organization.

Despite substantial progress in tensor data analysis, tensor methods remain relatively less well understood compared to their matrix and lower-order counterparts. The inherent multilinear structure of tensors introduces unique methodological and theoretical challenges that are not encountered in traditional settings. Moreover, while the statistical foundations of tensor methods have become robust, their adoption in real-world applications is still limited. As practical applications deepen, new challenges related to model specification, computational feasibility, and interpretability have gradually emerged. These inadequacies motivate us to develop statistically sound and computationally efficient methods for real tensor data analysis.

1.2 Contribution and Organization

In this thesis, we present four studies developing novel statistical methods with low-dimensional structures for different applications with tensor data. All the works

contribute to the community of tensor data analysis, from both methodological and applied perspectives. We present the works separately in four chapters.

In Chapter 2, we develop a *supervised tensor decomposition* (SupTD) method that incorporates multiple feature matrices as side information. While most tensor decomposition methods are developed for a single tensor observation, scientific studies often collect side information, in the form of node features and interactions thereof, together with the tensor data. Identifying the relationship between a high-dimensional tensor and side information is important yet challenging. Unlike unsupervised tensor decomposition, our supervised decomposition captures the effective dimension reduction of the data tensor confined to feature space of interest. An efficient alternating optimization algorithm with provable spectral initialization is further developed. Our proposal handles a broad range of data types, including continuous, count, and binary observations. We apply the method to diffusion tensor imaging data from human connectome project and multi-relational political network data. We identify the key global connectivity pattern and pinpoint the local regions that are associated with available features.

In Chapter 3, we consider the problem of multiway clustering in the presence of unknown degree heterogeneity. The allowance of degree heterogeneity provides great flexibility in clustering models, but the extra complexity poses significant challenges in both statistics and computation. Therefore, we develop a *degree-corrected tensor block model* (dTBM) with estimation accuracy guarantees. We present the phase transition of clustering performance based on the notion of angle separability, and we characterize three signal-to-noise regimes corresponding to different statistical-computational behaviors. In particular, we demonstrate that an intrinsic statistical-to-computational gap emerges only for tensors of order three or greater. Further, we develop an efficient polynomial-time algorithm that provably achieves exact clustering under mild signal conditions. The efficacy of our procedure is demonstrated through two data applications, one on human brain connectome project, and another on Peru Legislation network dataset.

Chapter 4 also considers the problem of multiway clustering, but with a special focus on the computational efficiency, as high computational costs hinder the

applications of tensor-based approaches to real-world large-scale data. We propose a *large-scale multiway clustering framework under tensor block model*, named LS-TBM, with accuracy guarantees. LS-TBM leverages seeded clustering to break down the expensive high-dimensional tensor clustering into two fast low-dimensional steps. Our two-step algorithm substantially reduces the time and space complexities from polynomial to logarithmic rates while maintaining the exact recovery of community structures, under certain signal conditions. We also establish the theoretical phase transition of LS-TBM performance with a key interplay between signal levels and seed sizes. Numerical experiments with synthetic data and real large-scale Uber Pickup data highlight LS-TBM's superior performance in practice.

In Chapter 5, our work addresses a key gap in understanding the mechanistic foundations for genotype-phenotype associations. While existing expression quantitative trait loci (eQTL) methods identify candidate loci affecting gene expression variants, they often neglect the crucial role of gene co-expression networks. We develop a network-based QTL framework, called *spectral network quantitative trait loci analysis* (snQTL), to map genetic loci affecting the gene co-expression network. Utilizing a tensor-based spectral approach, our snQTL method estimates the differential co-expression patterns and effectively identifies the associated genetic loci. Application of snQTL to three-spined sticklebacks revealed candidate loci missed by standard methods and estimated network changes led by genetic variants. Our result suggests the limitations of current approaches and highlights the potential of network-based functional loci discovery. Our method presents a novel solution for comparing biological networks, opening doors for deeper network analysis in genetics.

Last, we complete the thesis with conclusion remarks in Chapter 6.

1.3 Preliminaries

We now introduce several basic concepts and models used for statistical tensor data analysis. These concepts and models will be revisited in following works when we consider different tensor problems under specific context.

Tensor low-rankness

We start with tensor low-rankness, which are widely used to capture the low-dimensional structures of tensors. Unlike matrix rank, there are multiple definitions of tensor rank, corresponding to different tensor decompositions. In this thesis, we consider two popular low-rankness, CANDECOMP/PARAFAC (CP) low-rank and Tucker low-rank. Readers are referred to the survey Kolda and Bader (2009) for full introduction. Let $\mathcal{X} \in \mathbb{R}^{n_1 \times \cdots \times n_K}$ denote an order-K tensor.

We say $\mathcal{X}$ has the CP rank R if

$$\mathcal{X} = \sum_{r=1}^{R} \lambda^{(r)} \boldsymbol{a}_1^{(r)} \circ \cdots \circ \boldsymbol{a}_K^{(r)}, \tag{1.1}$$

where $\lambda^{(r)} > 0$ is the singular value, $\circ$ is the vector outer product such that $(\boldsymbol{a} \circ \boldsymbol{b} \circ \boldsymbol{c})_{i,j,k} = \boldsymbol{a}_i \boldsymbol{b}_j \boldsymbol{c}_k$, $\boldsymbol{a}_k^{(r)} \in \mathbb{R}^{n_k}$ for all $r = 1, \ldots, R$ and $k = 1, \ldots, K$. Based on specific context, people may add norm constrains $\|\boldsymbol{a}_i^{(r)}\|_2 = 1$, orthogonal constrains $\boldsymbol{a}_i^{(r)} \perp \boldsymbol{a}_i^{(r')}, r \neq r'$, and order constrain $\lambda^{(1)} \geq \ldots \geq \lambda^{(R)}$ for identifiability.

We say $\mathcal{X}$ has the Tucker rank $(R_1, \ldots, R_K)$ if

$$\mathcal{X} = \mathcal{S} \times_1 \boldsymbol{U}_1 \times_2 \cdots \times_K \boldsymbol{U}_K, \tag{1.2}$$

where $\mathcal{S} \in \mathbb{R}^{R_1 \times \cdots \times R_K}$ is a irreducible core tensor, $\boldsymbol{U}_k \in \mathbb{R}^{n_k \times R_k}$ are loading matrices presenting the subspaces on K modes. The tensor-by-matrix product operations $\times_k$ lead to the equivalent linear combinations

$$\mathcal{X}(i_1, \ldots, i_K) = \sum_{r_1=1}^{R_1} \cdots \sum_{i_K=1}^{R_K} \mathcal{S}(r_1, \ldots, r_K) \boldsymbol{U}_1(i_1, r_1) \ldots \boldsymbol{U}_K(i_K, r_K).$$

Tensor block structures

Tensor block structures can be considered as the higher-order extensions of matrix block structures. Block structure assumes that there exist community structures on multiple modes of the tensors, which is an alternative low-dimensional struc-

ture other than low-rankness. Here, we introduce the tensor block structures and the degree-corrected tensor block structures. Consider the order-K tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times \cdots \times n_K}$ and assume there exist $R_k \geq 1$ disjoint communities among n_k nodes for all $k = 1, \ldots, K$. Let functions $z_k : \{1, \ldots, n_K\} \mapsto \{1, \ldots, R_K\}$ represent the community assignments, where $z_k(i) = a$ indicating that the i-th node on the k-th mode belongs to the a-th community.

We say $\mathcal{X}$ has a tensor block structure if

$$\mathcal{X} = \mathcal{C} \times_1 \boldsymbol{M}_1 \times_2 \cdots \times_K \boldsymbol{M}_K, \tag{1.3}$$

where $\mathcal{C} \in \mathbb{R}^{R_1 \times \cdots \times R_K}$ is the core tensor collecting the block means among communities, $\boldsymbol{M}_k \in \{0, 1\}^{n_k \times R_k}$ are the membership matrices representing the community assignments such that $\boldsymbol{M}_k(i, a) = \mathbb{1}\{z_k(i) = a\}$. The block structure (1.3) can be considered as a special case of the Tucker decomposition (1.2), while $\mathcal{X}$ with $(R_1, \ldots, R_K)$ blocks dose not necessarily have the Tucker rank $(R_1, \ldots, R_K)$ as $\mathcal{C}$ may be reducible (i.e., two slides in $\mathcal{C}$ may be the same up to different scales) in clustering.

Degree-corrected block structure allows the degree heterogeneity for nodes within the same community. Let $\boldsymbol{\theta}_k = (\theta_k(1), \ldots, \theta_k(n_k))^T$ denote the degree heterogeneity for n_k nodes on the k-th mode. We say $\mathcal{X}$ has a degree-corrected tensor block structure if

$$\mathcal{X} = \mathcal{C} \times_1 \boldsymbol{\Theta}_1 \boldsymbol{M}_1 \times_2 \cdots \times_K \boldsymbol{\Theta}_K \boldsymbol{M}_K, \tag{1.4}$$

where $\boldsymbol{\Theta}_k = \mathrm{diag}(\boldsymbol{\theta}_k) \in \mathbb{R}^{n_k \times n_k}$ are diagonal matrices representing the node degree heterogeneity. Though $\mathcal{X}$ still has $(R_1, \ldots, R_K)$ communities, the degree-corrected block structure (1.4) leads to more structural flexibilities and more methodological challenges compared with the block structure (1.3).

Signal-plus-noise model

Despite all the works in this thesis addressing different problems, we are able to summarize these statistical methods with a signal-plus-noise model:

$$\mathcal{Y} = \mathcal{X} + \mathcal{E},$$

where $\mathcal{Y}$ is the tensor observation, $\mathcal{X}$ is the signal tensor with some low-dimensional structures, and $\mathcal{E}$ is the noise tensor. Our goal is to recover the signal tensor $\mathcal{X}$ or some critical components in $\mathcal{X}$, from the noisy observation $\mathcal{Y}$.

The choice of low-dimensional structure in $\mathcal{X}$ depends on the problem context. Specifically, in Chapter 2, we assume that $\mathcal{X}$ has a modified Tucker structure (1.2) and aim to estimate the decomposition components; in Chapters 3 and 4, we assume that $\mathcal{X}$ has degree-corrected (1.4) and regular (1.3) tensor block structures, respectively, tackling the multiway clustering problem under different setups; and in Chapter 5, we approximate $\mathcal{Y}$ by the rank-1 CP structured signal $\mathcal{X}$ (1.1), in which the approximated tensor singular value λ is used for multiple network hypothesis testing.

The assumption on the noise tensor $\mathcal{E}$ also depends on the application context and impacts the method development. In Chapter 2, we assume that $\mathcal{E}$ has independent zero-mean entries but do not limit the distribution of the entries; in Chapters 3 and 4, we narrow down the assumption on the noise to have independent zero-mean normal entries for main theoretical results, while the extension for Bernoulli noisy entries is also discussed; and in Chapter 5, we relax the limitation on $\mathcal{E}$ without any distributional assumptions and adopt a nonparametric permutation-based approach to obtain the testing result.

2 SUPERVISED TENSOR DECOMPOSITION

This chapter presents the *supervised tensor decomposition* method. Full work *Generalized tensor decomposition with features on multiple modes* (Hu et al., 2022) is published in *Journal of Computational and Graphical Statistics*.

2.1 Introduction

Multi-dimensional arrays, known as tensors, are often collected with side information on multiple modes in modern scientific and engineering studies. A popular example is in neuroimaging (Zhou et al., 2013). The brain connectivity networks are collected from a sample of individuals, accompanied by individual characteristics such as age, gender, and diseases status (see Figure 2.1a). Another example is in network analysis (Berthet and Baldin, 2020; Hoff, 2005). A typical social network consists of nodes that represent people and edges that represent friendships. Side information such as people's demographic information and friendship types are often available. In both examples, we are interested in identifying the variation in the data tensor (e.g., brain connectivities, social community patterns) that is affected by available features. These seemingly different scenarios pose a common yet challenging problem for tensor data modeling.

In addition to the aforementioned challenges, many tensor datasets consist of non-Gaussian measurements. Examples include the political interaction dataset (Nickel et al., 2011) which measures action counts between countries under various relations, and the brain connectivity network dataset (Zhou et al., 2013; Wang and Li, 2020b) which is a collection of binary adjacency matrices. Classical tensor decomposition methods are based on minimizing the Frobenius norm of deviation, leading to suboptimal predictions for binary- or count-valued response variables. A number of supervised tensor methods have been proposed (Lock and Li, 2018; Li and Zhang, 2017; Sun and Li, 2017; Hao et al., 2021; Raskutti et al., 2019) to address the tensor regression problem in various forms, such as scalar-to-tensor regression

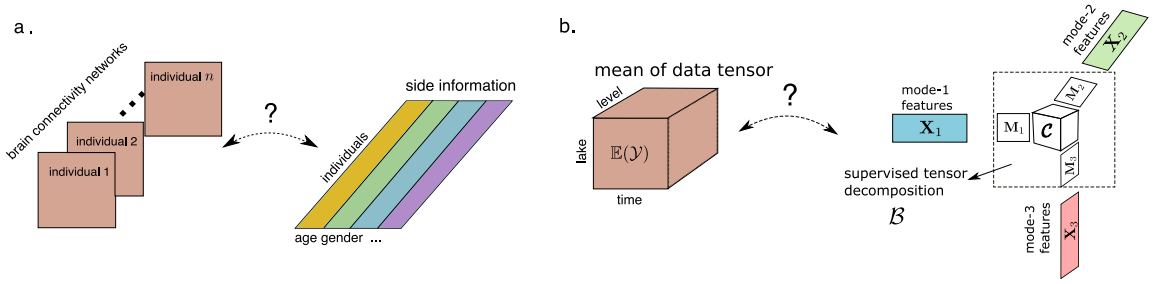

Figure 2.1: Examples of supervised tensor decomposition with side information. (a) Network population model. (b) Spatio-temporal growth model.

and tensor-response regression. These methods often assume Gaussian distribution for the tensor entries, or impose random designs for the feature matrices, both of which are less suitable for applications of our interest. The gap between theory and practice means a great opportunity to model paradigms and better capture the complexity in tensor data.

We present a general model and associated method for decomposing a data tensor whose entries are from exponential family with side information. We formulate the learning task as a structured regression problem, with tensor observation serving as the response, and the multiple side information as features. Figure 2.1b illustrates our model in the special case of order-3 tensors. A low-rank structure is imposed to the conditional mean of tensor observation, where unlike classical decomposition, the tensor factors $\boldsymbol{X}_k\boldsymbol{M}_k \in \mathbb{R}^{d_k \times r_k}$ belong to the space spanned by features $\boldsymbol{X}_k \in \mathbb{R}^{d_k \times p_k}$ for $k = 1, 2, 3$. The unknown matrices $\boldsymbol{M}_k \in \mathbb{R}^{p_k \times r_k}$ (referred to as "dimension reduction matrices") link the conditional mean to the feature spaces, thereby allowing the identification of variations in the tensor data attributable to the side information.

Our proposal blends the modeling power of generalized linear model (GLM) and the exploratory capability of tensor dimension reduction in order to take the best out of both worlds. We leverage GLM to allow heteroscedacity due to the mean-variance relationship in the non-Gaussian data. This flexibility is important in practice. Furthermore, our low-rank model on the (transformed) conditional

mean tensor effectively mitigates the curse of high dimensionality. In classical GLM, the sample size and feature dimension are well defined; however, in the tensor data analysis, we observe only one realization of an order-K tensor and up to K feature matrices. Both the number of tensor entries and feature dimension grow exponentially in K. Dimension reduction is therefore crucial for prediction and interpretability. We establish the statistical and algorithmic convergences of our estimator, and we quantify the gain in accuracy through simulations and case studies.

Our work is closely related to but also clearly distinctive from several lines of previous work. The first line is a class of *unsupervised* tensor decomposition such as classical Tucker and CP decomposition (De Lathauwer et al., 2000; Kolda and Bader, 2009) and generalized decomposition for non-Gaussian data (Chi and Kolda, 2012; Tarzanagh and Michailidis, 2019; Hong et al., 2020; Li, 2020). Regardless of the implementation, the unsupervised methods aim to find the best low-rank representation of a data tensor alone. In contrast, our model is a *supervised* tensor learning, which aims to identify the association between a data tensor and multiple features. The low-rank factorization is determined jointly by the tensor data and feature matrices in our model.

The second line of work studies the tensor-to-tensor regression. This category is further divided into three scenarios, depending on whether tensor is treated as predictors (Zhou et al., 2013; Raskutti et al., 2019; Han et al., 2020), as responses (Li and Zhang, 2017; Sun and Li, 2017; Zhang et al., 2018; Lock and Li, 2018; Luo et al., 2018), or both (Lock, 2018; Gahrooei et al., 2020). As we show in Section 2.4, our supervised tensor decomposition falls into this general category, and we provide a *provable* solution in new settings that have broader practical significance. Earlier work in this vein (Lock, 2018; Lock and Li, 2018; Gahrooei et al., 2020; Li, 2020) focuses on algorithm development, but not on the statistical accuracy. Li and Zhang (2017) introduces an envelope-based approach to identify sufficient dimension reduction (Adragni and Cook, 2009), but its theory is restricted to Gaussian data with one-sided feature matrix only. Raskutti et al. (2019) establishes the statistical accuracy for convex relaxed maximum likelihood estimator (MLE) of tensor regression.

However, convex relaxation for tensor optimizations suffers from computational intractability and statistical sub-optimality. Recent work has demonstrated the success of non-convex approaches in various tensor problems (Sun and Li, 2017; Zhang et al., 2018; Raskutti et al., 2019; Han et al., 2020); we go step further by allowing multiple feature matrices with either fixed or random designs. In Sections 2.3, we show that incorporating multiple feature matrices substantially improves the statistical accuracy. We provide a detailed comparison in Section 2.4; see Table 2.2.

The third line of work uses side information for various tensor learning tasks, such as for completion (Song et al., 2019) and for recommendation system (Farias and Li, 2019). These methods also study tensors with side information, but they take data-mining approaches to penalize predictions that are distant from side information. One important difference is that their goal is prediction but not parameter estimation. The effects of features and their interactions are not estimated in these data-driven approaches. In contrast, our goal is interpretable prediction, and we estimate the low-rank decomposition using a model-based approach. The model-based approaches benefits the interpretability in prediction. In this regards, our method opens up new opportunities for tensor data analysis in a wider range of applications.

Notation

We use lower-case letters (e.g., a, b, c) for scalars and vectors, upper-case boldface letters (e.g., $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$) for matrices, and calligraphy letters (e.g., $\mathcal{A}, \mathcal{B}, \mathcal{C}$) for tensors of order three or greater. We use $\boldsymbol{I}$ to denote the identity matrix whose dimension may vary from line by line given the contexts. Let $\mathcal{Y} = [\![y_{i_1,\ldots,i_K}]\!] \in \mathbb{R}^{d_1 \times \cdots \times d_K}$ denote an order-K $(d_1, \ldots, d_K)$-dimensional tensor, where K is the number of modes and also called the order. The multilinear multiplication of a tensor $\mathcal{Y} \in \mathbb{R}^{d_1 \times \cdots \times d_K}$ by matrices $\boldsymbol{X}_k = [\![x^{(k)}_{i_k,j_k}]\!] \in \mathbb{R}^{p_k \times d_k}$ is defined as

$$\mathcal{Y} \times_1 \boldsymbol{X}_1 \times \cdots \times_K \boldsymbol{X}_K = [\![\sum_{i_1,\ldots,i_K} y_{i_1,\ldots,i_K} x^{(1)}_{j_1,i_1} \cdots x^{(K)}_{j_K,i_K}]\!],$$

which results in an order-K $(p_1, \ldots, p_K)$-dimensional tensor. For ease of presentation, we use the shorthand $\mathcal{Y} \times \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K\}$ to denote the tensor-by-matrix product. For any two tensors $\mathcal{Y} = [\![y_{i_1,\ldots,i_K}]\!]$, $\mathcal{Y}' = [\![y'_{i_1,\ldots,i_K}]\!]$ of identical order and dimensions, their inner product is defined as

$$\langle \mathcal{Y},\ \mathcal{Y}' \rangle = \sum_{i_1,\ldots,i_K} y_{i_1,\ldots,i_K} y'_{i_1,\ldots,i_K}.$$

The tensor Frobenius norm and maximum norm are defined as

$$\|\mathcal{Y}\|_F = \langle \mathcal{Y},\ \mathcal{Y} \rangle^{1/2}, \quad \text{and} \quad \|\mathcal{Y}\|_\infty = \max_{i_1,\ldots,i_K} y_{i_1,\ldots,i_K}.$$

When a is a vector, we use $\|a\|_2 = \langle a, a \rangle^{1/2}$ to denote the vector 2-norm. We use $[d]$ to denote the d-set $[d] = \{1, \ldots, d\}$, and use $\mathbb{O}(d, r)$ to denote the collection of all d-by-r matrices with orthonormal columns; i.e., $\mathbb{O}(d, r) = \{\boldsymbol{P} \in \mathbb{R}^{d \times r} \colon \boldsymbol{P}^T \boldsymbol{P} = \boldsymbol{I}\}$.

A higher-order tensor can be reshaped into a lower-order object. We use $\text{vec}(\cdot)$ to denote the operation that reshapes the tensor into a vector, and $\text{Unfold}_k(\cdot)$ to denote the unfolding operation that reshapes the tensor along mode k into a matrix of size d_k-by-$\prod_{i \neq k} d_i$. We use $\text{rank}(\mathcal{Y}) = \boldsymbol{r}$ to denote the multilinear rank of an order-K tensor $\mathcal{Y}$, where $\boldsymbol{r} = (r_1, \ldots, r_K)$ is a length-K vector and r_k is the rank of matrix $\text{Unfold}_k(\mathcal{Y})$ for $k \in [K]$. For ease of notation, we allow the basic arithmetic operators (e.g., $+, -, \geq$) and univariate functions $f \colon \mathbb{R} \to \mathbb{R}$ to be applied to tensors in an element-wise manner. For two positive sequences $\{a_n\}$ and $\{b_n\}$, we use $a_n \lesssim b_n$ or $a_n = \mathcal{O}(b_n)$ to denote the fact that $a_n \leq C b_n$ for some constant $C > 0$.

2.2 Motivation and model

General framework for tensor decomposition

We begin with a general framework for supervised tensor decomposition and then discuss its implication in three concrete examples. Let $\mathcal{Y} = [\![y_{i_1,\ldots,i_K}]\!] \in \mathbb{R}^{d_1 \times \cdots \times d_K}$ denote an order-K data tensor. Suppose the side information is available on each

of the K modes. Let $\boldsymbol{X}_k = [\![x_{ij}]\!] \in \mathbb{R}^{d_k \times p_k}$ denote the feature matrix on the mode $k \in [K]$, where x_{ij} denotes the j-th feature value for the i-th tensor entity, for $(i, j) \in [d_k] \times [p_k]$.

We propose a multilinear conditional mean model between the data tensor and feature matrices. Assume that, conditional on the features $\boldsymbol{X}_k$, the entries of tensor $\mathcal{Y}$ are independent realizations from an exponential family distribution. Further, the conditional mean tensor admits the rank-$\boldsymbol{r}$ model with $\boldsymbol{r} = (r_1, \ldots, r_K)$,

$$\begin{aligned} \mathbb{E}(\mathcal{Y}|\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K) &= f\left(\mathcal{C} \times \{\boldsymbol{X}_1\boldsymbol{M}_1, \ldots, \boldsymbol{X}_K\boldsymbol{M}_K\}\right), \\ &\text{with } \boldsymbol{M}_k^T\boldsymbol{M}_k = \boldsymbol{I}_{r_k},\ \boldsymbol{M}_k \in \mathbb{R}^{p_k \times r_k} \quad \text{for all } k = 1, \ldots, K, \end{aligned} \tag{2.1}$$

where $\mathcal{C} \in \mathbb{R}^{r_1 \times \cdots \times r_K}$ is an unknown full-rank core tensor, $\boldsymbol{M}_k \in \mathbb{R}^{p_k \times r_k}$ are unknown factor matrices for all $k \in [K]$, $f(\cdot)$ is a known link function whose form depending on the data type of $\mathcal{Y}$, and $\times$ denotes the tensor-by-matrix product. The choice of link function is based on the assumed distribution family of tensor entries. Common choices of link functions include identity link for Gaussian distribution, logistic link for Bernoulli distribution, and exponential link for Poisson distribution. In general, dispersion parameters can also be included in the model. Because our main focus is the tensor decomposition under the mean model, we suppress the dispersion parameter in this section for ease of presentation.

Figure 2.1b provides a schematic illustration of our model. The features $\boldsymbol{X}_k$ affect the distribution of tensor entries in $\mathcal{Y}$ through the reduced features $\boldsymbol{X}_k\boldsymbol{M}_k$, which are r_k linear combinations of features on mode k. We call $\boldsymbol{M}_k$ the "dimension reduction matrix" or "tensor factors." The core tensor $\mathcal{C}$ collects the interaction effects between reduced features across K modes. We call $\mathcal{B} = \mathcal{C} \times \{\boldsymbol{M}_1, \ldots, \boldsymbol{M}_K\}$ the coefficient tensor, and $\Theta = \mathcal{B} \times \{\boldsymbol{X}_1, \cdots, \boldsymbol{X}_K\}$ the linear predictor. By the definition of multilinear rank, the model (2.1) implies the linear predictor Θ and coefficient tensor $\mathcal{B}$ are of rank-$\boldsymbol{r}$. The conditional mean tensor $\mathbb{E}(\mathcal{Y}|\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K)$ is however often high rank, due to the nonlinearity of the link function (Lee and Wang, 2021).

Our goal is to estimate the low-rank tensor $\mathcal{B}$, or equivalently, the core tensor

and factors $(\mathcal{C}, \boldsymbol{M}_1, \ldots, \boldsymbol{M}_K)$, from our model (2.1). We make several remarks about model identifiability. First, the identifiability of $\mathcal{B}$ requires the feature matrices $\boldsymbol{X}_k$ are of full column rank with $p_k \leq d_k$. We impose this rank non-deficiency assumption to $\boldsymbol{X}_k$; this is a mild condition common in literature (Lock and Li, 2018; Li and Zhang, 2017; Li, 2020). In the presence of rank deficiency, we recommend to remove redundant features from $\boldsymbol{X}_k$ before applying our method. Second, the decomposition $\mathcal{B} = \mathcal{C} \times \{\boldsymbol{M}_1, \ldots, \boldsymbol{M}_K\}$ are non-unique, as in standard tensor decomposition (Kolda and Bader, 2009). For any invertible matrices $\boldsymbol{O}_k \in \mathbb{R}^{r_k \times r_k}$, $\mathcal{B} = \mathcal{C} \times \{\boldsymbol{M}_1, \ldots, \boldsymbol{M}_K\} = \mathcal{C}' \times \{\boldsymbol{M}_1\boldsymbol{O}_1, \ldots, \boldsymbol{M}_K\boldsymbol{O}_K\}$ are two equivalent parameterizations with $\mathcal{C}' = \mathcal{C} \times \{\boldsymbol{O}_1^{-1}, \ldots, \boldsymbol{O}_K^{-1}\}$. To resolve this ambiguity, we impose orthonormality to $\boldsymbol{M}_k \in \mathbb{O}(p_k, r_k)$ and assess the estimation error of $\boldsymbol{M}_k$ using angle distance. The angle distance is invariant to orthogonal rotations due to its geometric definition. See Section 2.3 for more details. The orthonormality of $\boldsymbol{M}_k$ is imposed purely for technical convenience. This normalization incurs no impacts in our statistical inference, but may help with numerical stability in empirical optimization (De Lathauwer et al., 2000; Kolda and Bader, 2009). Finally, the problem size is quantified by p_k and d_k, where p_k specifies the number of features and d_k the number of samples at mode $k \in [K]$. Our theory treats the rank r_k as known and fixed, whereas both p_k and d_k are allowed to increase. The adaptation to unknown rank in practice will be addressed in Section 2.3.

Three examples

We give three seemingly different examples that can all be formulated as our supervised tensor decomposition model (2.1).

Example 2.1 (Spatio-temporal growth model)**.** *The growth curve model (Gabriel, 1998; Srivastava et al., 2008) was originally proposed as an example of bilinear model for matrix data, and we adopt its higher-order extension here. Let $\mathcal{Y} = [\![y_{ijk}]\!] \in \mathbb{R}^{d \times m \times n}$ denote the pH measurements of d lakes at m levels of depth and for n time points. Suppose the sampled lakes belong to q types, with p lakes in each type. Let $\{\ell_j\}_{j \in [m]}$ denote the sampled depth levels and $\{t_k\}_{k \in [n]}$ the time points. Assume that the expected pH trend in depth is a*

polynomial of order at most r and that the expected trend in time is a polynomial of order s. Then, the conditional mean model for the spatio-temporal growth can be represented as

$$\mathbb{E}(\mathcal{Y}|\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3) = \mathcal{C} \times \{\boldsymbol{X}_1\boldsymbol{M}_1,\ \boldsymbol{X}_2\boldsymbol{M}_2,\ \boldsymbol{X}_3\boldsymbol{M}_3\}, \tag{2.2}$$

where $\boldsymbol{X}_1 = \text{blockdiag}\{\mathbf{1}_p, \ldots, \mathbf{1}_p\} \in \{0,1\}^{d\times q}$ is the design matrix for lake types, and

$$\boldsymbol{X}_2 = \begin{pmatrix} 1 & \ell_1 & \cdots & \ell_1^r \\ 1 & \ell_2 & \cdots & \ell_2^r \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \ell_m & \cdots & \ell_m^r \end{pmatrix}, \quad \boldsymbol{X}_3 = \begin{pmatrix} 1 & t_1 & \cdots & t_1^s \\ 1 & t_2 & \cdots & t_2^s \\ \vdots & \vdots & \ddots & \vdots \\ 1 & t_n & \cdots & t_n^s \end{pmatrix}$$

are the design matrices for spatial and temporal effects, respectively, $\mathcal{C} \in \mathbb{R}^{r_1\times r_2\times r_3}$ is the unknown core tensor, and $\boldsymbol{M}_k$ are unknown dimension reduction matrices on each mode. The factors $\boldsymbol{X}_k\boldsymbol{M}_k$ are reduced features in the mean model (2.2)*. The spatial-temporal model is a special case of our supervised tensor decomposition model* (2.1)*, with features available on each of the three modes.*

Example 2.2 (Network population model)**.** *Network response model* (*Rabusseau and Kadri, 2016*) *is recently developed for neuroimaging analysis. The goal is to study the relationship between brain network connectivity pattern and features of individuals. Suppose we have a sample of n observations, $\{(\boldsymbol{Y}_i, \boldsymbol{x}_i) \colon i = 1, \ldots, n\}$, where for each individual $i \in [n]$, $\boldsymbol{Y}_i \in \{0,1\}^{d\times d}$ is the undirected adjacency matrix whose entries indicate presences/absences of connectivities between d brain nodes, and $\boldsymbol{x}_i \in \mathbb{R}^p$ is the individual's feature such as age, gender, cognition score, etc. The network-response model has the conditional mean*

$$\text{logit}(\mathbb{E}(\boldsymbol{Y}_i|\boldsymbol{x}_i)) = \mathcal{B} \times_3 \boldsymbol{x}_i, \quad \textit{for } i = 1, \ldots, n, \tag{2.3}$$

where $\mathcal{B} \in \mathbb{R}^{d\times d\times p}$ is a rank-(r_1, r_1, r_2) coefficient tensor, and $\mathcal{B}$ is assumed to be symmetric in the first two modes.

The model (2.3) *is a special case of our supervised tensor decomposition, with feature matrix on the last mode of the tensor. Specifically, we stack the network observations $\{\boldsymbol{Y}_i\}$*

together and obtain an order-3 response tensor $\mathcal{Y} \in \{0,1\}^{d\times d\times n}$*. Define a feature matrix* $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]^T \in \mathbb{R}^{n\times p}$*. Then, the model* (2.3) *has the equivalent representation of supervised tensor decomposition,*

$$\text{logit}(\mathbb{E}(\mathcal{Y}|\boldsymbol{X})) = \mathcal{C} \times \{\boldsymbol{M},\ \boldsymbol{M},\ \boldsymbol{X}\boldsymbol{M}'\},$$

where $\mathcal{C} \in \mathbb{R}^{r_1\times r_1\times r_2}$ *is the core tensor,* $\boldsymbol{M} \in \mathbb{R}^{d\times r_1}$ *is the dimension reduction matrix on the first two modes, and* $\boldsymbol{M}' \in \mathbb{R}^{p\times r_2}$ *is for the last mode.*

Example 2.3 (Dyadic data with node attributes)**.** *Dyadic dataset consists of measurements on pairs of objects. Common examples include graphs and networks. Let* $\mathcal{G} = (V, E)$ *denote a graph, where* $V = [d]$ *is the node set of the graph, and* $E \subset V \times V$ *is the edge set. Suppose that we also observe feature vector* $\boldsymbol{x}_i \in \mathbb{R}^p$ *associated to each node* $i \in V$*. A probabilistic model on the graph* $\mathcal{G} = (V, E)$ *can be described by the following matrix regression. The edge connects the two vertices* i *and* j *independently of other pairs, and the probability of connection is modeled as*

$$\text{logit}\left(\mathbb{P}\left((i,j) \in E\right)\right) = \boldsymbol{x}_i^T \boldsymbol{B} \boldsymbol{x}_j = \langle \boldsymbol{B},\ \boldsymbol{x}_i^T \boldsymbol{x}_j \rangle, \tag{2.4}$$

where $\boldsymbol{B} \in \mathbb{R}^{p\times p}$ *is a symmetric rank-*r *matrix. The low-rankness in* $\boldsymbol{B}$ *has demonstrated its success in modeling transitivity, balance, and communities in networks* (*Hoff, 2005*)*. We show that our supervised tensor decompostion* (2.1) *also incorporates the graph model as a special case. Let* $\mathcal{Y} = [\![y_{ij}]\!]$ *be a binary matrix where* $y_{ij} = \mathbb{1}_{(i,j)\in E}$*. Define* $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]^T \in \mathbb{R}^{n\times p}$*. Then, the graph model* (2.4) *can be expressed as*

$$\text{logit}(\mathbb{E}(\boldsymbol{Y}|\boldsymbol{X})) = \boldsymbol{C} \times \{\boldsymbol{X}\boldsymbol{M},\ \boldsymbol{X}\boldsymbol{M}\},$$

where $\boldsymbol{C} \in \mathbb{R}^{r\times r}$*,* $\boldsymbol{M} \in \mathbb{R}^{p\times r}$ *are from the singular value decomposition of* $\boldsymbol{B} = \boldsymbol{M}\boldsymbol{C}\boldsymbol{M}^T$*.*

In the above three examples and many other studies, researchers are interested in uncovering the variation in the data tensor that can be explained by features. Our supervised tensor decomposition (2.1) allows arbitrary numbers of feature matrices. When certain mode k has no side information, we set $\boldsymbol{X}_k = \boldsymbol{I}$ in the

model (2.1). In particular, our model (2.1) reduces to classical unsupervised tensor decomposition (De Lathauwer et al., 2000; Hong et al., 2020) when no side information is available; i.e., $\boldsymbol{X}_k = \boldsymbol{I}$ for all $k \in [K]$.

2.3 Estimation

Rank-constrained MLE

We develop a likelihood-based procedure to estimate $\mathcal{C}$ and $\boldsymbol{M}_k$ in (2.1). We adopt the exponential family as a flexible framework for different data types. In a classical generalized linear model with a scalar response y and feature $\boldsymbol{x}$, the density is expressed as

$$p(y|\boldsymbol{x}, \boldsymbol{\beta}) = c(y, \phi) \exp\left(\frac{y\theta - b(\theta)}{\phi}\right) \text{ with } \theta = \boldsymbol{\beta}^T \boldsymbol{x},$$

where $b(\cdot)$ is a known function, θ is the linear predictor, $\phi > 0$ is the dispersion parameter, and $c(\cdot)$ is a known normalizing function. The choice of link functions depends on the data types and on the observation domain of y, denoted $\mathbb{Y}$. For example, the observation domain is $\mathbb{Y} = \mathbb{R}$ for continuous data, $\mathbb{Y} = \mathbb{N}$ for count data, and $\mathbb{Y} = \{0, 1\}$ for binary data. The canonical link function f is chosen to be $f(\cdot) = b'(\cdot)$, the first-order derivative of $b(\cdot)$. Table 2.1 summarizes the canonical link functions for common types of distributions.

Data type	Gaussian	Poisson	Bernoulli
Domain $\mathbb{Y}$	$\mathbb{R}$	$\mathbb{N}$	$\{0, 1\}$
$b(\theta)$	$\theta^2/2$	$\exp(\theta)$	$\log(1 + \exp(\theta))$
link $f(\theta)$	θ	$\exp(\theta)$	$(1 + \exp(-\theta))^{-1}$

Table 2.1: Canonical links for common distributions.

In our context, we model the entries in data tensor $\mathcal{Y}$, conditional on linear predictor Θ, as independent draws from an exponential family. Ignoring constants that do not depend on Θ, the quasi log-likelihood of (2.1) is equal to Bregman

distance between $\mathcal{Y}$ and $b'(\Theta)$:

$$\mathcal{L}_{\mathcal{Y}}(\mathcal{C}, \boldsymbol{M}_1, \ldots, \boldsymbol{M}_K) = \langle \mathcal{Y}, \Theta \rangle - \sum_{i_1,\ldots,i_K} b(\theta_{i_1,\ldots,i_K}),$$
$$\text{where} \quad \Theta = \mathcal{C} \times \{\boldsymbol{X}_1\boldsymbol{M}_1, \ldots, \boldsymbol{X}_K\boldsymbol{M}_K\}. \tag{2.5}$$

We propose the constrained maximum quasi-likelihood estimate (MLE),

$$(\hat{\mathcal{C}}_{\text{MLE}}, \hat{\boldsymbol{M}}_{1,\text{MLE}}, \ldots, \hat{\boldsymbol{M}}_{K,\text{MLE}}) = \underset{(\mathcal{C},\boldsymbol{M}_1,\ldots,\boldsymbol{M}_K)\in\mathcal{P}(\boldsymbol{r})}{\arg\max} \mathcal{L}_{\mathcal{Y}}(\mathcal{C}, \boldsymbol{M}_1, \ldots, \boldsymbol{M}_K), \tag{2.6}$$

where the parameter space $\mathcal{P}(\boldsymbol{r})$ is defined by

$$\mathcal{P}(\boldsymbol{r}) = \left\{ (\mathcal{C}, \boldsymbol{M}_1, \ldots, \boldsymbol{M}_K) \;\middle|\; \boldsymbol{M}_k \in \mathbb{O}(p_k, r_k) \text{ for all } k \in [K],\ \|\Theta\|_\infty \le \alpha \right\}, \tag{2.7}$$

with a large constant $\alpha > 0$. Recall that $\mathcal{B} = \mathcal{C} \times \{\boldsymbol{X}_1, \ldots, \boldsymbol{M}_K\}$ by definition. Correspondingly, we estimate the coefficient tensor $\mathcal{B}$ by

$$\hat{\mathcal{B}}_{\text{MLE}} = \hat{\mathcal{C}}_{\text{MLE}} \times \{\hat{\boldsymbol{M}}_{1,\text{MLE}}, \ldots, \hat{\boldsymbol{M}}_{K,\text{MLE}}\}.$$

The maximum norm constraint on the linear predictor Θ is a technical condition to ensures the existence (boundedness) of MLE. The condition precludes the ill-defined MLE when the optimizer of (2.6) diverges to $\pm\infty$; this phenomenon may happen in logistic regression when the Bernoulli responses $\{0, 1\}$ are perfectly separable by covariates (Wang and Li, 2020b). For Gaussian models, no maximum norm constraint is needed. In Section 2.3, we show that setting α to an extremely large constant does not compromise the statistical rate in quantities of interest. In practice, the unbounded search is often indistinguishable from the bounded search, since the boundary constraint $\|\Theta\|_\infty \le \alpha$ would likely never be active. Similar techniques are commonly used in high-dimensional non-Gaussian problems (Wang and Li, 2020b; Han et al., 2020).

The optimization (2.6) is a non-convex problem with possibly local optimizers. We propose an alternating optimization algorithm to *approximately* solve (2.6). The

decision variables in the objective function (2.6) consist of $K+1$ blocks of variables, one for the core tensor $\mathcal{C}$ and K for the factor matrices $\boldsymbol{M}_k$. We notice that, if any K out of the $K+1$ blocks of variables are known, then the optimization reduces to a simple GLM with respect to the last block of variables. This observation leads to an iterative updating scheme for one block at a time while keeping others fixed. Given an initialization $(\hat{\mathcal{C}}^{(0)}, \hat{\boldsymbol{M}}_1^{(0)}, \ldots, \hat{\boldsymbol{M}}_K^{(0)})$ to be described in the next paragraph, the t-th iterate from the algorithm is denoted $(\hat{\mathcal{C}}^{(t)}, \hat{\boldsymbol{M}}_1^{(t)}, \ldots, \hat{\boldsymbol{M}}_K^{(t)})$ for $t = 1, 2, 3, \ldots$ The iteration scheme is detailed in Algorithm 1.

Algorithm 1 Supervised Tensor Decomposition with Side Information

Input: Response tensor $\mathcal{Y} \in \mathbb{R}^{d_1 \times \cdots \times d_K}$, feature matrices $\boldsymbol{X}_k \in \mathbb{R}^{d_k \times p_k}$ for $k = 1, \ldots, K$, target rank $\boldsymbol{r} = (r_1, \ldots, r_K)$, link function f, initialization $(\hat{\mathcal{C}}^{(0)}, \hat{\boldsymbol{M}}_1^{(0)}, \ldots, \hat{\boldsymbol{M}}_K^{(0)})$.

1: **for** $t = 1, 2, 3, \ldots$ **do**
2: **for** $k = 1$ to K **do**
3: Obtain the factor matrix $\hat{\boldsymbol{M}}_k^{(t)} \in \mathbb{R}^{p_k \times r_k}$ by a GLM with link function f.
4: Perform QR factorization $\hat{\boldsymbol{M}}_k^{(t)} = \boldsymbol{Q}_k \boldsymbol{R}_k$, where $\boldsymbol{Q}_k \in \mathbb{O}(p_k, r_k)$.
5: Update $\hat{\boldsymbol{M}}_k^{(t)} \leftarrow \boldsymbol{Q}_k$ and core tensor $\hat{\mathcal{C}}^{(t)} \leftarrow \hat{\mathcal{C}}^{(t)} \times_k \boldsymbol{R}_k$.
6: **end for**
7: Update the core tensor $\mathcal{C}$ by solving a GLM with $\mathrm{vec}(\mathcal{Y})$ as response, $\otimes_{k=1}^{K} [\boldsymbol{X}_k \boldsymbol{M}_k]$ as features, and f as link function. Here $\otimes$ denotes the Kronecker product of matrices.
8: **end for**

Output: factor estimate $(\hat{\mathcal{C}}^{(t)}, \hat{\boldsymbol{M}}_1^{(t)}, \ldots, \hat{\boldsymbol{M}}_K^{(t)})$ from the t-th iterate, and coefficient tensor estimate $\hat{\mathcal{B}}^{(t)} = \hat{\mathcal{C}}^{(t)} \times \{\hat{\boldsymbol{M}}_1^{(t)}, \ldots, \hat{\boldsymbol{M}}_K^{(t)}\}$.

We provide two initialization schemes, one with QR-adjusted spectral initialization (warm initialization), and the other with random initialization (cold initialization). The warm initialization is an extension of unsupervised spectral initialization (Zhang and Xia, 2018) to supervised setting with multiple feature matrices. Specifically, we project normalized data tensor $\bar{\mathcal{Y}}$ to the normalized multilinear feature space and obtain an unconstrained coefficient tensor $\hat{\mathcal{B}}^{(0)}$. We perform a rank-$\boldsymbol{r}$ higher-order SVD (HOSVD) on $\bar{\mathcal{B}}$, which yields the rank-constrained $\hat{\mathcal{B}}^{(0)}$.

The desired initialization is obtained by re-normalizing $\hat{\mathcal{B}}^{(0)}$ back to the original scales of features. The initialization scheme is described in Algorithm 2.

The warm initialization enjoys provable accuracy guarantees at a cost of extra technical assumptions (see Section 2.3). The cold initialization, on the other hand, shows robust in practice but its theoretical guarantee remains an open challenge (Luo and Zhang, 2021). We incorporate both options in our software package to provide flexibility to practitioners.

Algorithm 2 QR-adjusted spectral initialization

Input: Response tensor $\mathcal{Y} \in \mathbb{R}^{d_1 \times \cdots \times d_K}$, feature matrices $\boldsymbol{X}_k \in \mathbb{R}^{d_k \times p_k}$, Tucker rank $\boldsymbol{r}$.

1: Normalize date tensor $\bar{\mathcal{Y}} \leftarrow \mathcal{Y}$ for Gaussian model, $\bar{\mathcal{Y}} \leftarrow 2\mathcal{Y} - 1$ for Bernoulli model, and $\bar{\mathcal{Y}} \leftarrow \log(\mathcal{Y} + 0.5)$ for Poisson model.
2: Normalize feature matrices via QR factorization $\boldsymbol{X}_k = \boldsymbol{Q}_k \boldsymbol{R}_k$ for all $k \in [K]$.
3: Obtain $\bar{\mathcal{B}} \leftarrow \bar{\mathcal{Y}} \times \{\boldsymbol{Q}_1^T, \ldots, \boldsymbol{Q}_K^T\}$ by projecting $\bar{\mathcal{Y}}$ to the multilinear feature space.
4: Obtain $\hat{\mathcal{B}}^{(0)} \leftarrow \text{HOSVD}(\bar{\mathcal{B}}, \boldsymbol{r})$.
5: Normalize representation $\{\hat{\mathcal{C}}^{(0)}, \hat{\boldsymbol{M}}_1^{(0)}, \ldots, \hat{\boldsymbol{M}}_K^{(0)}\}$ such that $\hat{\mathcal{C}}^{(0)} \times \{\hat{\boldsymbol{M}}_1^{(0)}, \ldots, \hat{\boldsymbol{M}}_K^{(1)}\} = \hat{\mathcal{B}}^{(0)} \times \{\boldsymbol{R}_1^{-1}, \ldots, \boldsymbol{R}_K^{-1}\}$ and $\hat{\boldsymbol{M}}_k^{(0)} \in \mathbb{O}(p, r)$ for all $k \in [K]$.

Output: Core tensor $\hat{\mathcal{C}}^{(0)}$ and factors $\hat{\boldsymbol{M}}_k^{(0)}$ for all $k \in [K]$.

Statistical accuracy

This section presents the accuracy guarantees for both global and local optimizers of (2.6). We first provide the statistical accuracy for the global MLE (2.6). Then, we provide the convergence rate for the local optimizer from Algorithm 1 with warm initialization. The rate reveals an interesting interplay between statistical and computational efficiency. We show that a polynomial number of iterations suffices to reach the desired accuracy under certain assumptions. The empirical performance for cold initialization is also investigated.

For cleaner exposition, we present the results for balanced setting in this section, i.e., $p_1 = \cdots = p_K = p$, $r_1 = \cdots = r_K = r$, and $d_1 = \cdots = d_K = d$. The general

setting follows exactly the same framework and incurs only notational complexity. We are particularly interested in the high-dimensional regime in which both d and p grows while $p \leq d$. The requirement $p \leq d$ is necessary to ensure rank non-deficiency of feature matrices $\boldsymbol{X}_k$. The classical MLE theory is not directly applicable, because the number of unknown parameters grows with the size of data tensor. We leverage the recent development in random tensor theory and high-dimensional statistics to establish the error bounds of the estimation.

Assumption 2.1. *We make the following assumptions:*

A1. There exist two positive constants $c_1, c_2 > 0$ such that $c_1 \leq \sigma_{\min}(\boldsymbol{X}_k) \leq \sigma_{\max}(\boldsymbol{X}_k) \leq c_2$ for all $k \in [K]$. Here $\sigma_{\min}(\cdot)$ and $\sigma_{\max}(\cdot)$ denote the smallest and largest matrix singular values.

A1′. The feature matrices $\boldsymbol{X}_k$ are Gaussian designs with i.i.d. $N(0,1)$ entries.

A2. There exist two positive constants $L, U > 0$, such that $\min_{|\theta| \leq \alpha} b''(\theta) \geq \phi L$ and $\sup_{\theta \in \mathbb{R}} b''(\theta) \leq \phi U$. Here, α is the upper bound of the linear predictor in (2.6), and $b''(\cdot)$ denotes the second-order derivative.

The assumptions are fairly mild. Assumptions A1 and A1′ consider two separate scenarios about feature matrices. Assumption A1 is applicable when feature matrix is asymptotically non-singular and has bounded spectral norm, whereas Assumption A1′ imposes the commonly-used Gaussian design (Raskutti et al., 2019). The Assumption 2 is essentially imposed to the response variance because of the identity $\text{Var}(y|\theta) = \phi b''(\theta)$ (McCullagh and Nelder, 1989). The lower bound ensures the non-degeneracy of the variance in the feasible domain of θ, whereas the upper bound ensures the finiteness of the variance in the entire family. In fact, except for Poisson responses, most members in the exponential family, e.g., Gaussian, Bernoulli, and binomial responses, satisfy this condition.

Statistical accuracy for global optimizers

We need some extra notation to state the results in full generality. Recall that the factor matrices $\boldsymbol{M}_k$ are identifiable only up to orthogonal rotations. Therefore, we choose to use angle distance to assess the estimation accuracy of $\boldsymbol{M}_k$. For any two column-orthonormal matrices $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{O}(d, r)$ of same dimension, the angle distance is defined as

$$\sin\Theta(\boldsymbol{A}, \boldsymbol{B}) = \max\left\{ \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\|_2 \|\boldsymbol{y}\|_2} : \ \boldsymbol{x} \in \mathrm{Span}(\boldsymbol{A}), \ \boldsymbol{y} \in \mathrm{Span}(\boldsymbol{B}^\perp) \right\},$$

where $\mathrm{Span}(\cdot)$ represents the column space of the matrix. We use the superscript "true" to denote the true parameters from generic decision variables in optimization. For instance, $\mathcal{B}_{\text{true}}$ denotes the true coefficient tensor, whereas $\mathcal{B}$ denotes a decision variable in (2.5).

Define the signal level λ as the minimal singular value of the unfolded matrices obtained from $\mathcal{B}_{\text{true}}$,

$$\lambda = \min_{k \in [K]} \sigma_r(\mathrm{Unfold}_k(\mathcal{B}_{\text{true}})).$$

Intuitively, λ quantifies the level of rank non-degeneracy for the true coefficient tensor $\mathcal{B}_{\text{true}}$.

Theorem 2.2 (Statistical rate for global optimizers)**.** *Consider generalized tensor models with multiple feature matrices. Under Assumptions A1 and A2 with scaled feature matrices* $\bar{\boldsymbol{X}}_k = \sqrt{d}\boldsymbol{X}_k$*, or Assumptions A1′ and A2 with original feature matrices, we have*

$$\max_{k \in [K]} \sin^2 \Theta(\boldsymbol{M}_{k,\text{true}}, \hat{\boldsymbol{M}}_{k,MLE}) \lesssim \frac{\phi(r^K + Kpr)}{\lambda^2 d^K}, \quad \|\mathcal{B}_{true} - \hat{\mathcal{B}}_{MLE}\|_F^2 \lesssim \frac{\phi(r^K + Kpr)}{d^K}, \tag{2.8}$$

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

Theorem 2.2 establishes the statistical convergence for the global MLE (2.6). The result in (2.8) implies that the estimation has a convergence rate $\mathcal{O}(Kp/d^K)$ as $(p, d) \to \infty$. This rate agrees with intuition, since in our setting, the number of parameters with K feature matrices is of order $\mathcal{O}(Kp)$, whereas the number of

tensor entries $\mathcal{O}(d^K)$ corresponds to the total sample size. Because $p \leq d$, our rate is faster than $\mathcal{O}(d^{-(K-1)})$ obtained by tensor decomposition without features (Wang and Li, 2020b).

Inspection of our proof (Supplementary Notes) shows that the desired convergence rate holds not only for the MLE, but also for all local optimizers satisfying $\mathcal{L}_{\mathcal{Y}}(\mathcal{C}, \boldsymbol{M}_1, \ldots, \boldsymbol{M}_K) \geq \mathcal{L}_{\mathcal{Y}}(\mathcal{C}_{\text{true}}, \boldsymbol{M}_{1,\text{true}}, \ldots, \boldsymbol{M}_{K,\text{true}})$. The observation indicates the global optimality is not necessarily a serious concern in our context, as long as the convergent objective is large enough. In next section, we will provide the statistical accuracy for *local* optimizer with provable convergence guarantee, at a cost of extra signal requirement.

Empirical accuracy for local optimizers

The optimization (2.6) is a non-convex problem due to the low-rank constraint in the feasible set $\mathcal{P}$. Under mild conditions, our warm initialization enjoys stable performance, and the subsequent iterations further improve the accuracy via linear convergence; i.e. sequence of iterates generated by Algorithm 1 converges to optimal solutions at a linear rate.

Proposition 2.4 (Polynomial-time angle estimation)**.** *Consider Gaussian tensor models with $b(\theta) = \theta^2/2$ in the objective function* (2.5)*. Suppose the signal-to-noise ratio $\lambda^2/\phi \geq Cp^{K/2}d^{-K}$ for some sufficiently large universal constant $C > 0$. Under Assumption A1 with scaled feature matrices $\bar{\boldsymbol{X}}_k = \sqrt{d}\boldsymbol{X}_k$, or Assumption A1' with original feature matrices, the outputs from initialization Algorithm 2 and iteration Algorithm 1 satisfy the following two properties.*

(*a*) *With probability at least $1 - \exp(-p)$.*

$$\max_{k\in[K]} \sin^2 \Theta(\boldsymbol{M}_{k,\text{true}}, \hat{\boldsymbol{M}}_k^{(0)}) \leq \frac{1}{4}. \tag{2.9}$$

(*b*) *Let $t = 1, 2, 3, \ldots,$ denote the iteration. There exists a contraction parameter $\rho \in$*

$(0,1)$, *such that, with probability at least* $1-\exp(-p)$,

$$\max_{k\in[K]} \sin^2\Theta(\boldsymbol{M}_{k,\text{true}}, \hat{\boldsymbol{M}}_k^{(t)}) \lesssim \underbrace{\frac{\phi p}{\lambda^2 d^K}}_{\textit{statistical error}} + \underbrace{\rho^t \max_{k\in[K]} \sin\Theta^2(\boldsymbol{M}_{k,\text{true}}, \hat{\boldsymbol{M}}_k^{(0)})}_{\textit{algorithmic error}}. \quad (2.10)$$

Proposition 2.4 provides the estimation errors for algorithm outputs at initialization and at each of the subsequent iterations. The initialization bound (2.9) demonstrates the stability of warm initialization under a mild SNR requirement $\lambda^2/\phi \gtrsim p^{K/2}d^{-K}$. We can think of d as the sample size while p the number of parameters at mode K. This threshold is less stringent than $d^{K/2}$ required for unsupervised tensor decomposition features (Han et al., 2020; Zhang and Xia, 2018). The condition confirms that a higher sample size mitigates the required signal level. The iteration bound (2.10) consists of two terms: the first term is the statistical error, and the second is the algorithmic error. The algorithmic error decays exponentially with the number of iterations, whereas the statistical error remains the same as t grows. The statistical error is unavoidable and also appears in the global MLE; see Theorem 2.2.

As a direct consequence, we find the optimal iteration t after which the algorithmic error is negligible compared to statistical error.

Theorem 2.3 (Statistical rate for local optimizers)**.** *Consider the same condition as in Proposition 2.4 and the outputs by combining algorithms 1 and 2. There exists a constant* $C>0$*, such that, after* $t \gtrsim K\log_{1/\rho} p$ *iterations, our algorithm outputs satisfies*

$$\max_{k\in[K]} \sin^2\Theta(\boldsymbol{M}_{k,\text{true}}, \hat{\boldsymbol{M}}_k^{(t)}) \lesssim \frac{\phi p}{\lambda^2 d^K}, \quad \|\mathcal{B}_{\text{true}} - \hat{\mathcal{B}}^{(t)}\|_F^2 \lesssim \frac{\phi(r^K + Kpr)}{d^K}.$$

In practice, the signal level λ is unknown, so the assumption in Theorem 2.3 is challenging to verify in practice. We supply the theory by providing an alternative scheme – random initialization – and investigate its empirical performance. Figure 2.2 shows the trajectories of objective function for order-3 tensors based on model (2.1), where $d \in \{25, 30\}$, $p = 0.4d$, $r \in \{3, 6\}$ at all three modes. We consider

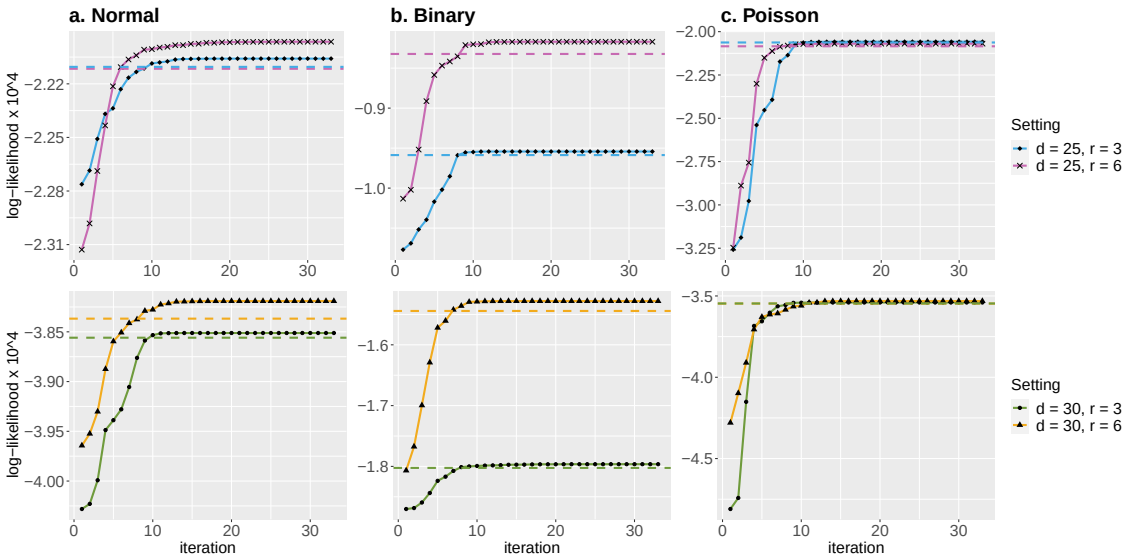

Figure 2.2: Trajectory of the objective function with various dimension d and rank r under (a) Gaussian (b) Bernoulli (c) Poisson models. The dashed line represents the objective value at true parameters.

data tensors with Gaussian, Bernoulli, and Poisson entries. Under all combinations of the dimension d, rank r, and type of the entries, Algorithm 1 converges quickly in a few iterations upon random initialization, and the objective values at convergent points are close to or larger than the value at true parameters. In the experiment we conduct, we find little difference in the final estimation errors between the two initialization schemes. Random initialization appears good enough for Algorithm 1 to find a convergent point with desired statistical guarantees. In practice, we recommend to run both warm and cold initializations, and choose the one with better convergent objective values.

We conclude this section by revisiting the three examples mentioned in Section 2.2.

Example 1 (Spatio-temporal growth model)**.** *The estimated type-by-time-by-space coefficient tensor converges at the rate* $\mathcal{O}\left((p+r+s)/(dmn)\right)$ *with* $(p, r, s) \leq (d, m, n)$.

The estimation achieves consistency as the dimension grows along either of the three modes.

Example 2 (Network population model)**.** *The estimated node-by-node-by-feature tensor converges at the rate* $\mathcal{O}\left((2d+p)/(d^2n)\right)$ *with* $p \leq n$*. The estimation achieves consistency as the number of individuals or the number of nodes grows.*

Example 3 (Dyadic data with node attributes)**.** *The estimated feature-by-feature matrix converges at the rate* $\mathcal{O}\left(p/d^2\right)$ *with* $p \leq d$*. Again, our estimation achieves consistency as the number of nodes grows.*

Rank selection and computational complexity

Our algorithm assumes the rank $\boldsymbol{r}$ is given. In practice, the rank is often unknown and must be determined from the data. We propose to use Bayesian information criterion (BIC) and choose the rank that minimizes BIC, where

$$\text{BIC}(\boldsymbol{r}) = -2\mathcal{L}_{\mathcal{Y}}(\hat{\mathcal{C}}, \hat{\boldsymbol{M}}_1, \ldots, \hat{\boldsymbol{M}}_K) + p_e(\boldsymbol{r}) \log(\textstyle\prod_k d_k). \tag{2.11}$$

Here, $p_e(\boldsymbol{r}) \stackrel{\text{def}}{=} \sum_k (p_k - r_k) r_k + \prod_k r_k$ is the effective number of parameters in the model. We choose $\hat{\boldsymbol{r}}$ that minimizes $\text{BIC}(\boldsymbol{r})$ via grid search. Our choice of BIC aims to balance between the goodness-of-fit for the data and the degree of freedom in the population model. We evaluate the empirical performance of BIC in Section 2.5.

The computational complexity of our Algorithm is $\mathcal{O}\left(d \sum_k p_k^3\right)$ for each iteration, where $d = \prod_k d_k$ is the total size of the data tensor. The update of K factor matrices is $\mathcal{O}(d \sum_k r_k^3 p_k^3)$ via standard GLM routines. Furthermore, we demonstrate that, under certain SNR conditions, a polynomial number of iterations suffices to reach the desired statistical accuracy. Therefore, the total computational cost is polynomial in p and d.

2.4 Connection to other tensor regression methods

We compare our supervised tensor decomposition (**SupTD**) with recent 12 tensor methods in the literature. Table 2.2 summarizes these methods with their prop-

erties from four aspects: i) model specification, ii) number of feature matrices allowed, (iii) capability of addressing non-Gaussian response, and (iv) capability of addressing non-independent noise. The four closet methods to our are **SupCP** (Lock and Li, 2018), **Envelope** (Li and Zhang, 2017), **mRRR** (Luo et al., 2018) and **GLSNet** (Zhang et al., 2018); these methods all relate a data tensor to feature matrices with low-rank structure on the coefficients. As seen from the table, our method is the only one that allows multiple feature matrices among the five. **Envelope** and **SupCP** are developed for Gaussian data, and the Gaussianity facilities flexible extension to non-independent noise. In particular, **Envelope** allows noise correlation in Kronecker structured form, whereas **SupCP** allows noise correlation implicitly through decomposing the latent factors into fixed effects (related to features) and random effects (unrelated to features). On the other hand, the other three methods (**mRRR**, **GLSNet** and **SupTD**) are developed for exponential family distribution with possibly non-additive noise. The generality makes the full modeling of noise correlation computationally challenging. We will compare the numerical performance of these methods in Section 2.5.

Our model also has a close connection to higher-order interaction model (Hao et al., 2020) and tensor-to-tensor regression (Lock, 2018). Model (2.1) can be viewed as a regression model with across-mode interactions in the reduced feature space. We take an order-3 tensor under the Gaussian model for illustration. Let $\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{W}$ denote the feature matrix on mode $k = 1, 2, 3$, respectively. Suppose that each mode has two-dimensional reduced features, denoted $\boldsymbol{M}_1\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2]$, $\boldsymbol{M}_2\boldsymbol{Z} = [\boldsymbol{z}_1, \boldsymbol{z}_2]$, $\boldsymbol{M}_3\boldsymbol{W} = [\boldsymbol{w}_1, \boldsymbol{w}_2]$. Here $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{w}_1, \boldsymbol{w}_2$ are column vectors. Then the model (2.1) is equivalent to a regression model with across-mode interactions

$$\mathbb{E}(y_{ijk}|\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{W}) = c_{111}\boldsymbol{x}_{1i}\boldsymbol{z}_{1j}\boldsymbol{w}_{1k} + c_{121}\boldsymbol{x}_{i1}\boldsymbol{z}_{j2}\boldsymbol{w}_{k1} + \cdots + c_{221}\boldsymbol{x}_{2i}\boldsymbol{z}_{2j}\boldsymbol{w}_{1k} + c_{222}\boldsymbol{x}_{2i}\boldsymbol{z}_{2j}\boldsymbol{w}_{2k},$$

where $[\![c_{ijk}]\!] \in \mathbb{R}^{2\times 2\times 3}$ are unknown interaction effects, $\boldsymbol{x}_{1i}$ denotes the i-th entry in the feature vector $\boldsymbol{x}_1$, and similar notations apply to other features. Note that lower-order interactions are naturally incorporated if we include an intercept column in the reduced feature matrices. The above example shows the connection of our

Method	Model	No. of features	non-Gaussianity	Non-independence
SupTD (Ours)	$\mathbb{E}\mathcal{Y} = f(\mathcal{B} \times \{\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3\}),\ \mathcal{B} = \mathcal{C} \times \{\boldsymbol{M}_1, \boldsymbol{M}_2, \boldsymbol{M}_3\}$	3	$\surd$	$\times$
GCP, CP-ARP, CORALS	$\mathbb{E}\mathcal{Y} = f(\llbracket \boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3 \rrbracket)$	0	$\surd$	$\times$
DCOT	$\mathbb{E}\mathcal{Y} = f((\mathcal{C}_1 + \mathcal{C}_2) \times \{\boldsymbol{M}_1, \boldsymbol{M}_2, \boldsymbol{M}_3\})$	0	$\surd$	$\times$
LRT, CRT	$y_n = \langle \mathcal{B},\ \mathcal{X}_n \rangle + \epsilon_n$, various structure on $\mathcal{B}$	0	$\times$	$\times$
STAR	$y_n = \sum_m \langle \mathcal{B}_m, \mathcal{F}_m(\mathcal{X}_{ijk}) \rangle + \epsilon_n$, sparse-CP $\mathcal{B}_m$	0	$\times$	$\times$
SupCP	$\mathcal{Y} = \llbracket \boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3 \rrbracket + \mathcal{E},\ \boldsymbol{A}_1 = \boldsymbol{X}\boldsymbol{B} + \mathcal{E}',\ \mathcal{E} \perp \mathcal{E}'$	1	$\times$	$\surd$
mRRR	$\mathbb{E}\boldsymbol{Y} = f(\boldsymbol{X}\boldsymbol{B})$, low-rank $\boldsymbol{B}$	1	$\surd$	$\times$
Envelope	$\mathcal{Y} = \mathcal{B} \times_3 \boldsymbol{X} + \mathcal{E},\ \mathcal{B} = \mathcal{C} \times \{\boldsymbol{M}_1, \boldsymbol{M}_2, \boldsymbol{I}\},\ \mathcal{E} \sim \mathcal{TN}(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \boldsymbol{I})$	1	$\times$	$\surd$
GLSNet	$\mathbb{E}\mathcal{Y} = f(\mathbf{1} \otimes \boldsymbol{\Theta} + \mathcal{B} \times_3 \boldsymbol{X})$, low-rank $\boldsymbol{\Theta}$, sparse $\mathcal{B}$	1	$\surd$	$\times$
STORE	$\mathcal{Y} = \mathcal{B} \times_3 \boldsymbol{X} + \mathcal{E}$, sparse-CP $\mathcal{B}$	1	$\times$	$\times$

Table 2.2: Comparison of tensor regression/factorization methods. We focus on order-3 tensors for illustration. Calligraphic letters denote tensors, bold capital letters denote matrices, and little letters denote scalars. The dimension of tensors and matrices can be identified from the contexts.

- Data: tensor response $\mathcal{Y}$, feature matrices $\boldsymbol{X}, \boldsymbol{X}_k$, predictor tensor $\mathcal{X}_n$, scalar response y_n, sample index n, tensor mode $k = 1, 2, 3$.
- Parameter: Tuckor factors $\boldsymbol{M}_k$, CP factors $\boldsymbol{A}_k$, CP decomposition $\llbracket \boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3 \rrbracket$, coefficient tensor and matrix $\mathcal{B}, \mathcal{B}_m, \boldsymbol{\Theta}, \boldsymbol{B}$.
- Function: a known link function $f(\cdot)$, a known basis function $\mathcal{F}_m(\cdot)$.
- Noise: Gaussian tensor with i.i.d. entries $\mathcal{E}, \mathcal{E}'$, Gaussian tensor with Kronecker covariance $\mathcal{E} \sim \mathcal{TN}(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \boldsymbol{I})$, meaning $\text{Cov}(\text{vec}(\mathcal{E})) = \boldsymbol{\Sigma}_1 \otimes \boldsymbol{\Sigma}_2 \otimes \boldsymbol{I}$.

- GCP: Generalized canonical polyadic tensor decomposition (Hong et al., 2020);
- CP-APR: CP alternating Poisson regression (Chi and Kolda, 2012);
- CORALS: Generalized co-clustering method (Li, 2020);
- DCOT: Double core tensor decomposition (Tarzanagh and Michailidis, 2019);
- SupCP: Supervised PARAFAC/CANDECOMP factorization (Lock and Li, 2018);
- mRRR: Mixed-response reduced-rank regression (Luo et al., 2018);
- Envelope: Parsimonious tensor response regression (Li and Zhang, 2017);
- GLSNet: Generalized connectivity matrix response regression (Zhang et al., 2018);
- STORE: Sparse tensor response regression (Sun and Li, 2017);
- LTR: Low-rank tensor regression (Han et al., 2020);
- CRT: Convex regularized multi-response tensor regression (Raskutti et al., 2019);
- STAR: Sparse tensor additive regression (Hao et al., 2021).

supervised tensor decomposition to multivariate regressions.

2.5 Numerical experiments

We evaluate the empirical performance of our supervised tensor decomposition (**SupTD**) through simulations. We consider order-3 tensors with a range of distribution types. Unless otherwise specified, the conditional mean tensor is generated form model (2.1), where the core tensor entries are i.i.d. drawn from Uniform[-1,1], the factor matrix $\boldsymbol{M}_k$ is uniformly sampled with respect to Haar measure from matrices with orthonormal columns. The feature matrix $\boldsymbol{X}_k$ is either an identity matrix (i.e., no feature available) or Gaussian random matrix with i.i.d. entries from $N(0,1)$. The linear predictor $\Theta = \mathcal{C} \times \{\boldsymbol{M}_1\boldsymbol{X}_1, \boldsymbol{M}_2\boldsymbol{X}_2, \boldsymbol{M}_3\boldsymbol{X}_3\}$ is scaled such that $\|\Theta\|_\infty = 1$. Conditional on the linear predictor $\Theta = [\![\theta_{ijk}]\!]$, the entries in the tensor $\mathcal{Y} = [\![y_{ijk}]\!]$ are drawn independently according to three probabilistic models:

(a) Gaussian model: continuous tensor entries $y_{ijk} \sim N(\alpha\theta_{ijk}, 1)$.

(b) Poisson model: count tensor entries $y_{ijk} \sim \text{Poisson}\left(e^{\alpha\theta_{ijk}}\right)$.

(c) Bernoulli model: binary tensor entries $y_{ijk} \sim \text{Bernoulli}\left(\frac{e^{\alpha\theta_{ijk}}}{1+e^{\alpha\theta_{ijk}}}\right)$.

Here $\alpha > 0$ controls the magnitude of the effect size, which is also the maximum norm of coefficient tensor as in (2.7). In each experiment, we report the summary statistics averaged across 30 simulation replications.

Finite-sample performance

The first experiment assesses the selection accuracy of our BIC criterion (2.11). We consider the balanced situation where $d_k = d$, $p_k = 0.4d_k$ for $k = 1, 2, 3$. We set $\alpha = 4$ and consider various combinations of dimension d and rank $\boldsymbol{r} = (r_1, r_2, r_3)$. For each combination, we minimize BIC using a grid search from $(r_1 - 3, r_2 - 3, r_3 - 3)$ to $(r_1 + 3, r_2 + 3, r_3 + 3)$. We remove invalid rank such as $r_{\max}^2 \geq \prod_{k=1}^3 r_k$ and use parallel search to reduce the computational cost. Table 2.3 reports the selected

True rank $\boldsymbol{r}$	$d = 20$ (Gaussian)	$d = 40$ (Gaussian)	$d = 20$ (Poisson)	$d = 40$ (Poisson)
$(3,\ 3,\ 3)$	$(\mathbf{3.0},\ \mathbf{3.0},\ \mathbf{3.0})$	$(\mathbf{3.0},\ \mathbf{3.0},\ \mathbf{3.0})$	$(\mathbf{3.0},\ \mathbf{3.0},\ \mathbf{3.0})$	$(\mathbf{3.0},\ \mathbf{3.0},\ \mathbf{3.0})$
$(4,\ 4,\ 6)$	$(3.0,\ 3.0,\ \mathbf{4.6})$	$(\mathbf{4.0},\ \mathbf{4.0},\ \mathbf{5.3})$	$(3.0,\ 3.0,\ \mathbf{5.3})$	$(\mathbf{4.0},\ \mathbf{4.0},\ \mathbf{5.6})$
$(6,\ 8,\ 8)$	$(5.0,\ 5.0,\ 5.0)$	$(\mathbf{6.0},\ \mathbf{8.0},\ \mathbf{8.0})$	$(5.0,\ 5.0,\ 6.7)$	$(\mathbf{6.0},\ \mathbf{8.0},\ \mathbf{8.0})$

Table 2.3: Rank selection via BIC. The estimated ranks are averaged across 30 simulation. Bold number indicates the ground truth is within two standard deviations of the estimate.

rank averaged over $n_{\text{sim}} = 30$ replicates. We find that in the high-rank setting with $d = 20$, the selected rank slightly underestimates the true rank, and the accuracy immediately improves when either the dimension increases to $d = 40$ or the rank reduces to $\boldsymbol{r} = (3, 3, 3)$. This agrees with our expectation, because in the tensor decomposition, the sample size is related to the number of tensor entries. A larger d implies a larger sample size, so the BIC selection becomes more accurate.

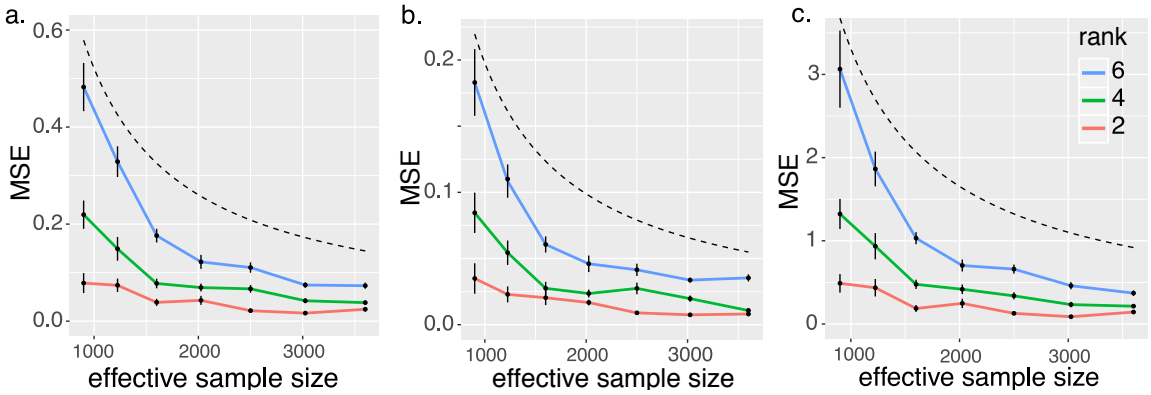

Figure 2.3: Estimation error against effective sample size. The three panels plot the MSE when the response tensors are generated from (i) Gaussian (ii) Poisson and (iii) Bernoulli models. The dashed curves correspond to $\mathcal{O}(1/d^2)$.

The second experiment evaluates the accuracy when features are available on all modes. We set $\alpha = 10, d_k = d, p_k = 0.4d_k, r_k = r \in \{2, 4, 6\}$ and increase d from 30 to 60. Our theoretical analysis suggests that $\hat{\mathcal{B}}$ has a convergence rate $\mathcal{O}(d^{-2})$ in this setting. Figure 2.3 plots the mean squared error (MSE) $\|\hat{\mathcal{B}} - \mathcal{B}_{\text{true}}\|_F^2$ versus

the effective sample size d^2 under three different distribution models. We find that the empirical MSE decreases roughly at the rate of $1/d^2$, which is consistent with our theoretical results. We also observe that, tensors with higher rank tend to yield higher estimation errors, as reflected by the upward shift of the curves as r increases. Indeed, a larger r implies a higher model complexity and thus greater difficulty in the estimation.

Comparison with other tensor methods

We compare our supervised tensor decomposition with three other tensor methods:

- Supervised PARAFAC/CANDECOMP factorization (**SupCP**, (Lock and Li, 2018)).
- Parsimonious tensor response regression (**Envelope**, (Li and Zhang, 2017));
- Mixed-response reduced-rank regression (**mRRR**, (Luo et al., 2018));
- Generalized connectivity matrix response regression (**GLSNet**, (Zhang et al., 2018));

These four methods are the closest methods to ours, in that they all relate a data tensor to feature matrices with low-rank structure on the coefficients. We consider Gaussian and Bernoulli tensors in experiments. For methods not applicable for Bernoulli data (**SupCP** and **Envelope**), we provide the algorithm $\{-1, 1\}$-valued tensors as inputs. Because **mRRR** allows matrix response only, we provide the algorithm the unfolded matrix of response tensor as inputs. We measure the accuracy using the response error defined as $1 - \text{Cor}(\hat{\mathcal{Y}}, f(\Theta_{\text{true}}))$, where $\hat{\mathcal{Y}}$ is the fitted tensor from each method, and $f(\Theta_{\text{true}})$ is the true conditional mean of the tensor. Note that the response error is a scale-insensitive metric; a smaller error implies a better fit of the model.

The comparison is assessed from three aspects: (i) benefit of incorporating features from multiple modes; (ii) prediction error with respect to sample size;

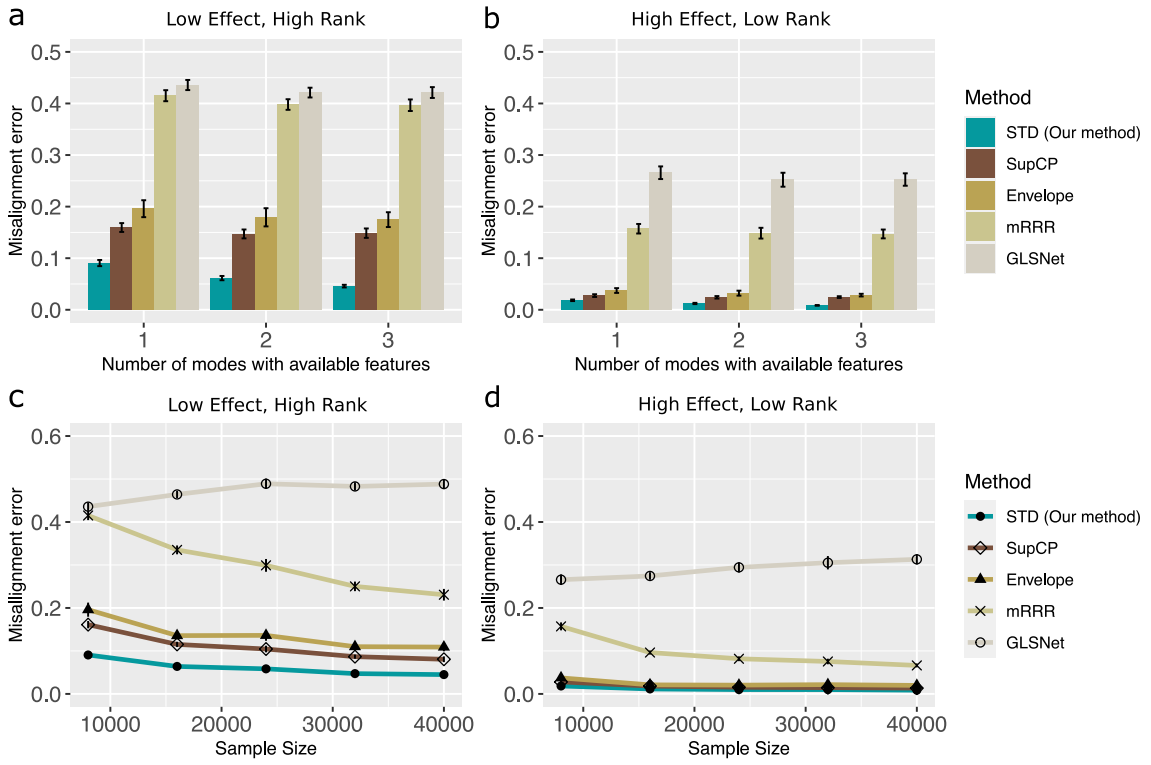

Figure 2.4: Comparison between tensor methods with Gaussian data. Panels (a) and (b) plot estimation error versus the number of modes with available features. Panels (c) and (d) plot ME versus the effective sample size d^2. We consider rank $\boldsymbol{r} = (3, 3, 3)$ (low), $\boldsymbol{r} = (4, 5, 6)$ (high), and effect size $\alpha = 3$ (low), $\alpha = 6$ (high).

(iii) robustness of model misspecification. We use similar simulation setups as in our first experiment in last section. We consider rank $\boldsymbol{r} = (3, 3, 3)$ (low) vs. $(4, 5, 6)$ (high), effect size $\alpha = 3$ (low) vs. 6 (high), dimension d ranging from 20 to 100 for modes with features, and $d = 20$ for modes without features. The method **Envelope** and **mRRR** require the tensor rank as inputs, respectively. For fairness, we provide both algorithms the true rank. The methods **SupCP** and **GLSNet** use different notions of model rank, and **GLSNet** takes sparsity as an input. We use a grid search to set the hyperparameters in **SupCP** and **GLSNet** that give the best performance.

Figure 2.4a-b shows the impact of features to estimation error. We see that our **SupTD** outperforms others, especially in the low-effect high-rank setting. As the number of informative modes increases, the **SupTD** exhibits a reduction in error whereas others remain unchanged. The accuracy gain in Figure 2.4 demonstrates the benefit of incorporating informative features from multiple modes. In addition, we find that the relative performance among the competing methods reveals the benefits of low-rankness. The second best method is **SupCP** which imposes low-rankness on three modes; the next one is **Envelope** which imposes low-rankness on two modes; the less favorable one is **mRRR** which imposes low-rank structure on one mode only; the worst one is **GLSNet** which imposes sparsity but no low-rankness on the feature effects.

Figure 2.4c-d compares the prediction error with respect to effective sample size d^2. For fair comparison, we consider the setting with feature matrix on one mode only. We find that our **SupTD** method has similar performance as **Envelope** and **SupCP** in the high-effect low-rank regime, whereas the improvement becomes more pronounced in the low-effect high-rank regime. The latter setting is notably harder, and our **SupTD** method shows advantage in addressing this challenge. Among other methods, **Envelope**, **SupCP**, and **mRRR** show decreasing errors as d increases, implying the benefits of low-rankness methods. In contrast, **GLSNet** suffers from non-decreasing error and indicates the poor fit of sparsity methods in addressing low-rank data.

We also evaluate the performance comparison with Bernoulli tensors. Figure 2.5 indicates the necessarity of generalized model in addressing non-Gaussian data. Indeed, methods that assume Gaussiannity (**Envelope** and **SupCP**) perform less favorably in Bernoulli setting (Figure 2.5c) compared to Gaussian setting (Figure 2.4c). Our method shows improved accuracy as the number of informative features increases (Figure 2.5a-b). In the absence of multiple features, our method still performs favorably compared to others (Figure 2.5c-d), for the same reasons we have argued in Gaussian data.

Lastly, we assess the performance of our method **SupTD** under model misspecification. We consider two aspects: (i) non-independent noise, and (ii) sparse

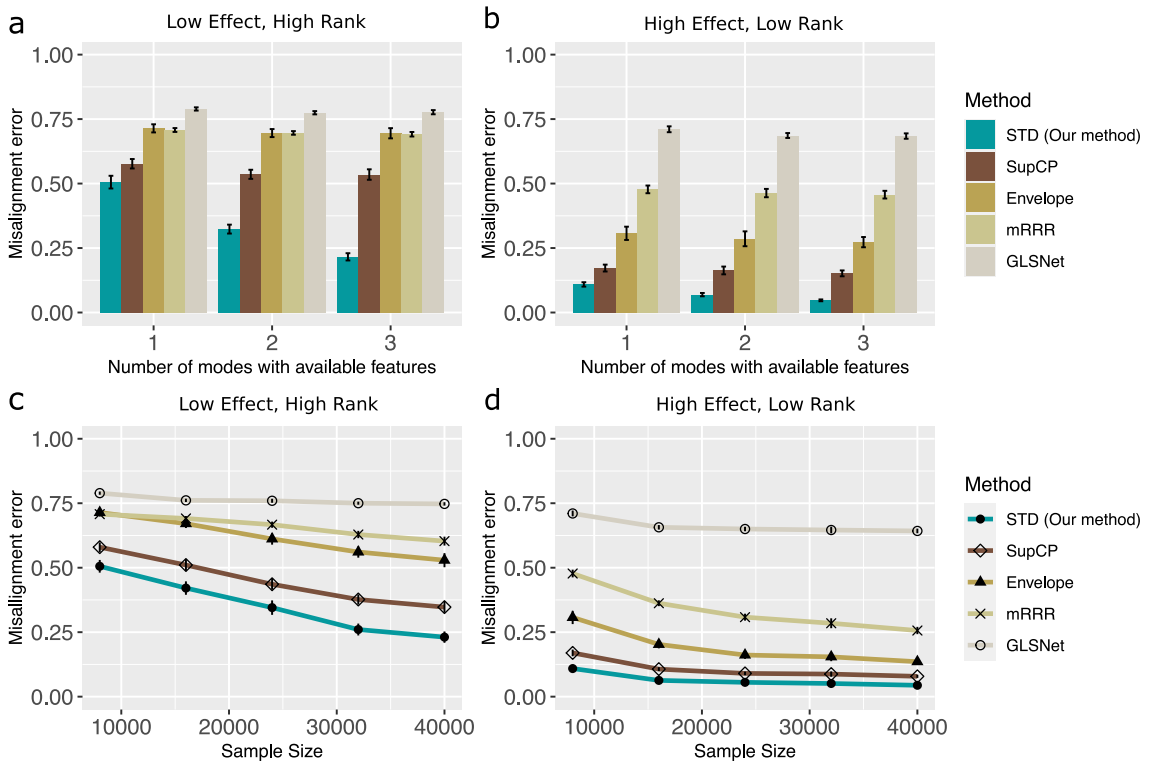

Figure 2.5: Comparison between tensor methods with Binary data. The panel legends are the same as in Figure 2.4.

feature effects. Note that our method **SupTD** imposes neither of these two assumptions, so the experiment allows us to assess the robustness. We select competing methods from Table 2.2 that specifically address these two aspects. We use **Envelope** and **SupCP** as benchmark for noise correlation experiment, and **GLSNet** for sparsity experiment.

Figure 2.6a-b assesses the impact of noise correlation to the estimation accuracy. The data is simulated from **Envelope** model with envelope dimensions $r = (3, 3)$ (low) and $(4, 5)$ (high). The noise is generated from a zero-mean Gaussian tensor with Kronecker structured covariance; see Supplementary Notes for details. As expected, **Envelope** shows the best performance in the high correlation setting. Remarkably, we find that our method **SupTD** has comparable and sometimes

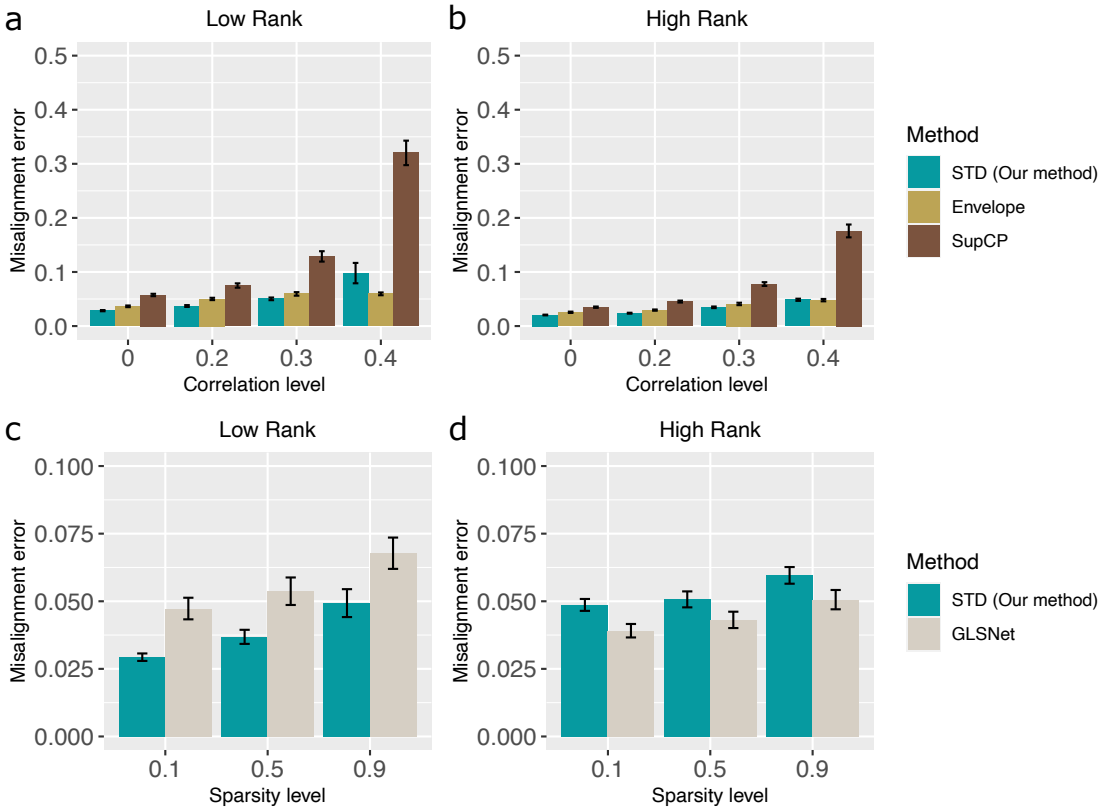

Figure 2.6: Comparison between tensor methods under model misspecification. Panels (a)-(b) assess the noise correlation, and panels (d)-(d) assess the sparsity.

better performance when noise correlation is moderate-to-low. In contrast, **SupCP** appears less suitable in this setting. Although **SupCP** allows noise correlation implicitly through latent random factors, the induced correlation may not belong to the Kronecker covariance structure in the simulation.

Figure 2.6c-d assesses the impact of sparsity to estimation performance. We generate data from **GLSNet** model, except that we modify the coefficient tensor to be joint sparse and low-rank (the original **GLSNet** model assumes full-rankness on the coefficient tensor). The sparsity level (x-axis in Figure 2.6c-d) quantifies the proportion of zero entries in the coefficient tensor. Since neither our method **SupTD**

nor **GLSNet** follow the simulated model, this setting allows a fair comparison. We find that our method outperforms **GLSNet** in the low-rank setting, whereas **GLSNet** shows a better performance in the high-rank setting. This observation suggests the robustness of our method to sparsity when the tensor of interest is simultaneously low-rank and sparse. When sparsity is the only salient structure, then methods specifically addressing sparsity would provide a better fit.

2.6 Data applications

We apply our supervised tensor decomposition to two datasets. The first application studies the brain networks in response to individual attributes (i.e., feature on one mode), and the second application focuses on multi-relational network analysis with dyadic attributes (i.e., features on two modes).

Application to human brain connection data

The Human Connectome Project (HCP) aims to build a network map that characterizes the anatomical and functional connectivity within healthy human brains (Van Essen et al., 2013). We follow the preprocessing procedure as in (Desikan et al., 2006) and parcellate the brain into 68 regions of interest. The dataset consists of 136 brain structural networks, one for each individual. Each brain network is represented as a 68-by-68 binary matrix, where the entries encode the presence or absence of fiber connections between the 68 brain regions. We consider four individual features: gender (65 females vs. 71 males), age 22-25 ($n = 35$), age 26-30 ($n = 58$), and age 31+ ($n = 43$). The preprocessed dataset is released in our R package `tensorregress`. The goal is to identify the connection edges that are affected by individual features. A key challenge in brain network is that the edges are correlated; for example, the nodes in edges may be from a same brain region, and it is of importance to take into account the within-dyad dependence.

We perform the supervised tensor decomposition to the HCP data. The BIC selection suggests a rank $\boldsymbol{r} = (10, 10, 4)$ with quasi log-likelihood $\mathcal{L}_y = -174654.7$.

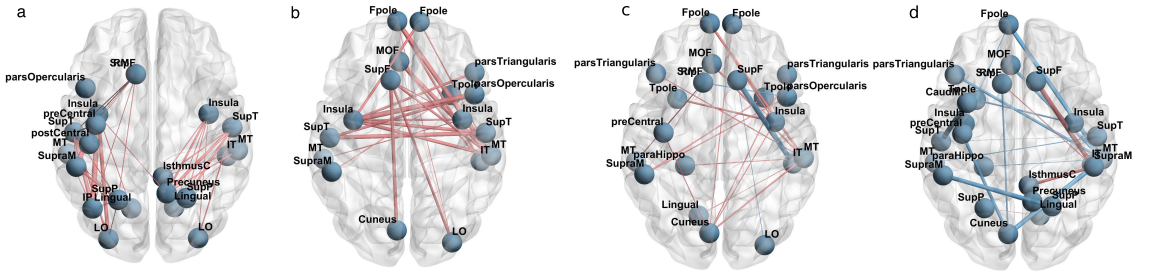

Figure 2.7: Top edges with large effects. (a) Global effect; (b) Female effect; (c) Age 22-25; (d) Age 31+. Red edges represent positive effects and blue edges represent negative effects. The edge-width is proportional to the magnitude of the effect size.

We utilize the sum-to-zero contrasts in coding the feature effects, and depict only the top 3% edges whose connections are non-constant across the sample. Figure 2.7 shows the top edges with high effect size, overlaid on the Desikan atlas brain template (Desikan et al., 2006). We find that the global connection exhibits clear spatial separation, and that the nodes within each hemisphere are more densely connected with each other (Figure 2.7a). In particular, the superior-temproal (*SupT*), middle-temporal (*MT*) and Insula are the top three popular nodes in the network. Interestingly, female brains display higher inter-hemispheric connectivity, especially in the frontal, parental and temporal lobes (Figure 2.7b). This is in agreement with a recent study showing that female brains are optimized for inter-hemispheric communication (Ingalhalikar et al., 2014). We find several edges with declined connection in the group Age 31+. Those edges involve Frontal-pole (*Fploe*), superior-frontal (*SupF*) and Cuneus nodes. The Frontal-pole region is known for its importance in memory and cognition, and the detected decline with age further highlights its biological importance.

Application to political relation data

The second application studies the multi-relational networks with node-level attributes. We consider *Nations* dataset (Nickel et al., 2011) which records 56 relations among 14 countries between 1950 and 1965. The multi-relational networks can be

organized into a $14 \times 14 \times 56$ binary tensor, with each entry indicating the presence or absence of an action, such as "sending tourist to", "export", "import", between countries. The 56 relations span the fields of politics, economics, military, religion, etc. In addition, country-level attributes are also available, and we focus on the following six features: *constitutional, catholics, law ngo, political leadership, geography*, and *medicine ngo*. The goal is to identify the variation in connections due to country-level attributes and their interactions.

We apply our tensor model to the *Nations* data. The multi-relational network $\mathcal{Y}$ is a binary data tensor, and the country attributes $\boldsymbol{X} \in \mathbb{R}^{14\times 6}$ are features on both the 1st and 2nd modes. We use BIC as guidance to select the rank of coefficient tensor $\mathcal{B}$. Since several rank configurations give similar BIC values, we present here the most interpretable results with $\boldsymbol{r} = (4, 4, 4)$. Detailed rank selection procedure is in Supplementary Notes. We perform the supervised tensor decomposition and obtain the dimension reduction matrices $\hat{\boldsymbol{M}}_k$ from the model. Then we apply K-mean clustering to dimension reduction matrix on each of the modes. Table S6 in Supplementary Notes shows the K-means clustering of the 56 relations based on the dimension reduction matrix on the 3rd mode. We find that the relations reflecting the similar aspects of actions are grouped together. In particular, Cluster I consists of military relations such as *violentactions*, *warnings* and *militaryactions*; Clusters II and III capture the economic relations such as *economicaid, booktranslations, tourism*; and Cluster IV represents the political alliance and territory relations.

To investigate the effects of dyadic attributes towards connections, we depict the estimated coefficients $\hat{\mathcal{B}} = [\![\hat{b}_{ijk}]\!]$ for several relation types (Figure 2.8). The entry $\hat{b}_{ijk}$ estimates the contribution, at the logit scale, of feature pair (i, j) (ith feature for the "sender" country and jth feature for the "receiver" country) towards the connection of relation k. Several interesting findings emerge from the observation. We find that relations belonging to a same cluster tend to have similar feature effects. For example, the relations "warning" and "violentactions" are classified into Cluster I, and both exhibit similar effect patterns (Figure 2.8a-b). Moreover, the feature *constitutional* has a strong effect in the relation "violentactions" and "warning", whereas the effect is weaker in the relation "treaties". The result is

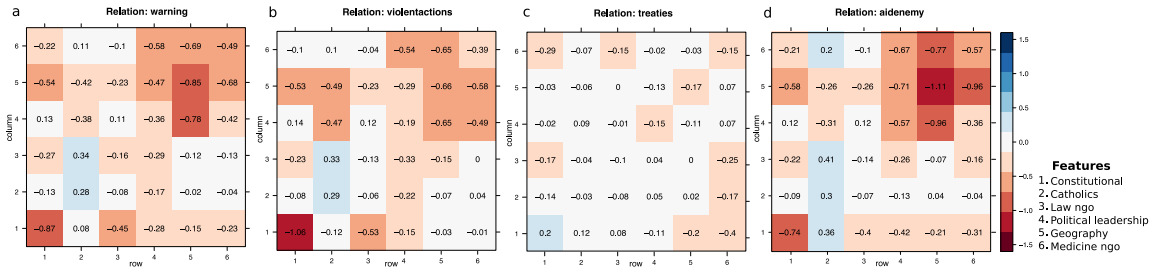

Figure 2.8: Estimated feature effects in the *Nations* data analysis. Panels (a)-(d) represent the estimated effects of country-level attributes towards the connection probability, for relations *warning*, *violentactions*, *treaties*, and *aidenemy*, respectively.

plausible because the constitutional attributes affect political actions more often than economical actions. The entries in $\mathcal{B}$ are useful for revealing interaction effects in a context-specific way. From Figure 2.8, we find a strong interaction between *geography* and *political leadership* in the relation "warning", and a strong interaction between *geogrphy* and *medicine ngo* in the relation "aidenemy". The relation-specific effect pattern showcases the applicability of our method in revealing complex interactions.

2.7 Discussion and future work

We have developed a supervised tensor decomposition method with side information on multiple modes. One important challenge of tensor data analysis is the complex interdependence among tensor entries and between multiple features. Our approach incorporates side information as feature matrices in the conditional mean tensor. The empirical results demonstrate the improved interpretability and accuracy over previous approaches. Applications to the brain connection and political relationship datasets yield conclusions with sensible interpretations, suggesting the practical utility of the proposed approach.

There are several possible extensions from the work. We have provided accuracy guarantees for parameter estimation in the supervised tensor model. Statistical

inference based on tensor decomposition is an important future direction. Measures of uncertainty, such as confidence envelope for space estimation, would be useful. One possible approach would be performing parametric bootstrap (Efron and Tibshirani, 1994) to assess the uncertainty in the estimation. For example, one can simulate tensors from the fitted low-rank model based on the estimates, and then assess the empirical distribution of the estimates. While being simple, bootstrap approach is often computationally expensive for large-scale data. Another possibility is to leverage recent development in debiased inference with distributional characterization (Chen et al., 2019b). This approach has led to fruitful results for matrix data analysis. Uncertainly quantification involving tensors are generally harder, and establishing distribution theory for tensor estimation remains an open problem.

One assumption made by our method is that tensor entries are conditionally independent given the linear predictor Θ. This assumption can be extended by introducing a more general mixed-effect tensor model. For example, in the special case of Gaussian model, we can model the first two moments of data tensor using

$$\begin{aligned}\mathbb{E}(\mathcal{Y}|\boldsymbol{X}_1,\ldots,\boldsymbol{X}_K) &= \mathcal{C}\times\{\boldsymbol{X}_1\boldsymbol{M}_1,\cdots,\boldsymbol{X}_K\boldsymbol{M}_K\},\\ \mathrm{Var}(\mathcal{Y}|\boldsymbol{X}_1,\ldots,\boldsymbol{X}_K) &= \boldsymbol{\Phi}_1\otimes\cdots\otimes\boldsymbol{\Phi}_K,\end{aligned}$$

where $\boldsymbol{\Phi}_k\in\mathbb{R}^{d_k\times d_k}$ is the unknown covariance matrix on the mode $k\in[K]$. For general exponential family, an additional mean-variance relationship should also be considered. The joint estimation of mean model Θ and variance model $\boldsymbol{\Phi}_k$ will lead to more efficient estimation in the presence of unmeasured confounding effects. However, the introduction of unknown covariance matrices $\boldsymbol{\Phi}_k$ dramatically increases the number of parameters in the problem. Suitable regularization such as graphical lasso or specially-structured covariance (Li and Zhang, 2017; Lock and Li, 2018) should be considered. The extension of tensor modeling with heterogeneous mixed-effects will be an important future direction.

Although we have presented the data applications in the context of order-3 data tensors, the framework of the supervised tensor decomposition applies to a

variety of multi-way datasets. One possible application is the integrative analysis of omics data, in which multiple types of omics measurements (gene expression, DNA methylation, microRNA) are collected in the same set of individuals (Lock et al., 2013; Wang et al., 2019). Other applications include time-series tensor data with multiple side information. Exploiting the benefits and properties of supervised tensor decomposition in specialized task will boost scientific discoveries.

3 MULTIWAY CLUSTERING – DEGREE-CORRECTED BLOCK MODEL

This chapter presents the *multiway clustering* method under the degree-corrected block model (dTBM). Full work *Multiway spherical clustering via degree-corrected tensor block models* (Hu and Wang, 2023) is published in *IEEE Transactions on Information Theory* and *International Conference on Artificial Intelligence and Statistics* (*AISTATS*).

3.1 Introduction

We study the problem of multiway clustering based on a data tensor. The goal of multiway clustering is to identify a checkerboard structure from a noisy data tensor. Figure 3.1 illustrates the noisy tensor and the underlying checkerboard structures discovered by multiway clustering methods. In the hypergraph example, the multiway clustering aims to identify the underlying block partition of nodes based on their higher-order connectivities; therefore, we also refer to the clustering as *higher-order clustering*. The most common model for higher-order clustering is called *tensor block model* (TBM) (Wang and Zeng, 2019), which extends the usual matrix stochastic block model (Abbe, 2018) to tensors. The matrix analysis tools, however, are sub-optimal for higher-order clustering. Developing tensor tools for solving block models has received increased interest recently (Wang and Zeng, 2019; Chi et al., 2020; Han et al., 2022a).

The classical tensor block model suffers from drawbacks to model real world data in spite of the popularity. The key underlying assumption of block model is that all nodes in the same community are exchangeable; i.e., the nodes have no individual-specific parameters apart from the community-specific parameters. However, the exchangeability assumption is often non-realistic. Each node may contribute to the data variation by its own multiplicative effect. We call the unequal node-specific effects the *degree heterogeneity*. Such degree heterogeneity appears commonly in social networks. Ignoring the degree heterogeneity may seriously

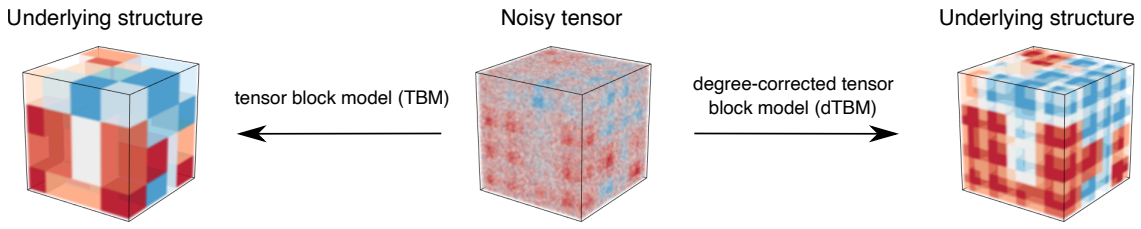

Figure 3.1: Examples for order-3 tensor block model (TBM) with and without degree correction. Both TBM and dTBM have four communities on each mode, while dTBM allows a richer structure with degree heterogeneity.

mislead the clustering results. For example, the regular block model fails to model the member affiliation in the Karate Club network (Bickel and Chen, 2009) without addressing degree heterogeneity.

The *degree-corrected tensor block model* (dTBM) has been proposed recently to account for the degree heterogeneity (Ke et al., 2019). The dTBM combines a higher-order checkerboard structure with degree parameter $\boldsymbol{\theta} = (\boldsymbol{\theta}(1), \ldots, \boldsymbol{\theta}(p))^T$ to allow heterogeneity among p nodes. Figure 3.1 compares the underlying structures of TBM and dTBM with the same number of communities. The dTBM allows varying values within the same community, thereby allowing a richer structure. To solve dTBM, we project clustering objects to a unit sphere and perform iterative clustering based on angle similarity. We refer to the algorithm as the *spherical* clustering; detailed procedures are in Section 3.4. The spherical clustering avoids the estimation of nuisance degree heterogeneity. The usage of angle similarity brings new challenges to the theoretical results, and we develop new polar-coordinate based techniques in the proofs.

Our contributions. The primary goal of this paper is to provide both statistical and computational guarantees for dTBM. Our main contributions are summarized below.

- We develop a general dTBM and establish the identifiability for the uniqueness of clustering using the notion of angle separability.

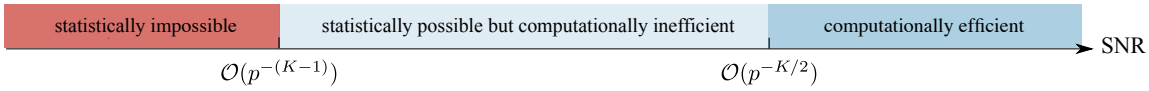

Figure 3.2: SNR thresholds for statistical and computational limits in order-K dTBM with dimension $(p, ..., p)$ and $K \geq 2$. The SNR gap between statistical possibility and computational efficiency exists only for tensors with $K \geq 3$.

- We present the phase transition of clustering performance with respect to three different statistical and computational behaviors. We characterize, for the first time, the critical signal-to-noise (SNR) thresholds in dTBMs, revealing the intrinsic distinctions among (vector) one-dimensional clustering, (matrix) biclustering, and (tensor) higher-order clustering. Specific SNR thresholds and algorithm behaviors are depicted in Figure 3.2.

- We provide an angle-based algorithm that achieves exact clustering *in polynomial time* under mild conditions. Simulation and data studies demonstrate that our algorithm outperforms existing higher-order clustering algorithms.

The last two contributions, to our best knowledge, are new to the literature of dTBMs.

Related work. Our work is closely related to but also distinct from several lines of existing research. Table 3.1 summarizes the most relevant models.

- *Block model for clustering.* The block model such as stochastic block model (SBM) and degree-corrected SBM has been widely used for matrix clustering problems. The theoretical properties and algorithm performance for matrix block models have been well-studied (Gao et al., 2018); see the review paper (Abbe, 2018) and the references therein. However, The tensor counterparts are relatively less understood.

- *Tensor block model.* The (non-degree) tensor block model (TBM) is a higher-order extension of SBM, and its statistical-computational properties are investigated in recent literatures (Wang and Zeng, 2019; Han et al., 2022a; Ghosh-

dastidar and Dukkipati, 2017a). Some works (Ahn et al., 2018) study the TBM with sparse observations, while, others (Wang and Zeng, 2019; Han et al., 2022a) and our work focus on the dense regime. Extending results from non-degree to degree-corrected model is highly challenging. Our dTBM parameter space is equipped with angle-based similarity and nuisance degree parameters. The extra complexity makes the Cartesian coordinates based analysis (Han et al., 2022a) non-applicable to our setting. Towards this goal, we have developed a new polar coordinates based analysis to control the model complexity. We have also developed a new angle-based iteration algorithm to achieve optimal clustering rates *without the need of estimating nuisance degree parameters*.

- *Degree-corrected block model.* The hypergraph degree-corrected block model (hDCBM) and its variant have been proposed in the literature (Ke et al., 2019; Yuan et al., 2022b). For this popular model, however, the optimal statistical-computational rates remain an open problem. Our main contribution is to provide a sharp statistical and computational critical phase transition in dTBM literature. In addition, our algorithm results in a faster *exponential* error rate, in contrast to the *polynomial* rate in Ke et al. (2019). The original hDCBM (Ke et al., 2019) is designed for binary observations only, and we extend the model to both continuous and binary observations. We believe our results are novel and helpful to the community. See Figure 3.2 for overview of our results.

- *Global-to-local algorithm strategy.* Our methods generalize the recent global-to-local strategy for matrix learning (Gao et al., 2018; Chi et al., 2019; Yun and Proutiere, 2016) to tensors (Han et al., 2022a; Ahn et al., 2018; Kim et al., 2018). Despite the conceptual similarity, we address several fundamental challenges associated with this non-convex, non-continuous problem. We show the insufficiency of the conventional tensor HOSVD (De Lathauwer et al., 2000), and we develop a weighted higher-order initialization that relaxes the singular-value gap separation condition. Furthermore, our local iteration leverages the angle-based clustering in order to avoid explicit estimation of

	Gao et al. (2018)	Ahn et al. (2018)	Han et al. (2022a)	Ghoshdastidar and Dukkipati (2017a)	Ke et al. (2019)	**Ours**
Allow tensors of arbitrary order	×	√	√	√	√	√
Allow degree heterogeneity	√	×	×	√	√	√
Singular-value gap-free clustering	√	√	√	×	×	√
Misclustering rate (for order K^*)	-	$p^{-(K-1)}\alpha^{-1}$**	$\exp(-p^{K/2})$	p^{-1}	p^{-2}	$\exp(-p^{K/2})$
Consider sparse observation	×	√	×	×	×	×

Table 3.1: Comparison between previous methods with our method. *We list the result for order-K tensors with $K \geq 3$ and general number of communities $r = \mathcal{O}(1)$. **The parameter $\alpha = f(p) > 0$ denotes the sparsity level which is some function of dimension p.

degree heterogeneity. Our bounds reveal the interesting interplay between the computational and statistical errors. We show that our final estimate *provably* achieves the exact clustering within only polynomial-time complexity.

Notation. We use lower-case letters (e.g., a, b) for scalars, lower-case boldface letters (e.g., $\boldsymbol{a}, \boldsymbol{\theta}$) for vectors, upper-case boldface letters (e.g., $\boldsymbol{X}, \boldsymbol{Y}$) for matrices, and calligraphy letters (e.g., $\mathcal{X}, \mathcal{Y}$) for tensors of order three or greater. We use $\mathbf{1}_p$ to denote a vector of length p with all entries to be 1. We use $|\cdot|$ for the cardinality of a set and $\mathbb{1}\{\cdot\}$ for the indicator function. For an integer $p \in \mathbb{N}_+$, we use the shorthand $[p] = \{1, 2, ..., p\}$. For a length-p vector $\boldsymbol{a}$, we use $a(i) \in \mathbb{R}$ to denote the i-th entry of $\boldsymbol{a}$, and use $\boldsymbol{a}_I$ to denote the sub-vector by restricting the indices in the set $I \subset [p]$. We use $\|\boldsymbol{a}\| = \sqrt{\sum_i a^2(i)}$ to denote the ℓ_2-norm, $\|\boldsymbol{a}\|_1 = \sum_i |a_i|$ to denote the ℓ_1 norm of $\boldsymbol{a}$. For two vector $\boldsymbol{a}, \boldsymbol{b}$ of the same dimension, we denote the angle between $\boldsymbol{a}, \boldsymbol{b}$ by

$$\cos(\boldsymbol{a}, \boldsymbol{b}) = \frac{\langle \boldsymbol{a}, \boldsymbol{b} \rangle}{\|\boldsymbol{a}\| \, \|\boldsymbol{b}\|},$$

where $\langle \boldsymbol{a}, \boldsymbol{b} \rangle$ is the inner product of two vectors and $\cos(\boldsymbol{a}, \boldsymbol{b}) \in [-1, 1]$. We make the convention that $\cos(\boldsymbol{a}, \boldsymbol{b}) = \cos\left(\boldsymbol{a}^T, \boldsymbol{b}^T\right)$.

Let $\mathcal{Y} \in \mathbb{R}^{p_1 \times \cdots \times p_K}$ be an order-K $(p_1, ..., p_K)$-dimensional tensor. We use $\mathcal{Y}(i_1, \ldots, i_K)$ to denote the $(i_1, \ldots, i_K)$-th entry of $\mathcal{Y}$. The multilinear multiplication of a tensor $\mathcal{S} \in \mathbb{R}^{r_1 \times \cdots \times r_K}$ by matrices $\boldsymbol{M}_k \in \mathbb{R}^{p_k \times r_k}$ results in an order-K $(p_1, \ldots, p_K)$-dimensional tensor $\mathcal{X}$, denoted

$$\mathcal{X} = \mathcal{S} \times_1 \boldsymbol{M}_1 \times \cdots \times_K \boldsymbol{M}_K,$$

where the entries of $\mathcal{X}$ are defined by

$$\mathcal{X}(i_1, \ldots, i_K) = \sum_{(j_1, \ldots, j_K)} \mathcal{S}(j_1, \ldots, j_K) \boldsymbol{M}_1(i_1, j_1) \cdots \boldsymbol{M}_K(i_K, j_K).$$

For a matrix $\boldsymbol{Y}$, we use $\boldsymbol{Y}_{i:}$ (respectively, $\boldsymbol{Y}_{:i}$) to denote the i-th row (respectively, i-th column) of the matrix. Similarly, for an order-3 tensor, we use $\mathcal{Y}_{::i}$ to denote the i-th matrix slide of the tensor. We use $\text{Ave}(\cdot)$ to denote the operation of taking averages across elements and $\text{Mat}_k(\cdot)$ to denote the unfolding operation that reshapes the tensor along mode k into a matrix. For a symmetric tensor $\mathcal{X} \in \mathbb{R}^{p \times \cdots \times p}$, we omit the subscript and use $\text{Mat}(\mathcal{X}) \in \mathbb{R}^{p \times p^{K-1}}$ to denote the unfolding. For two sequences $\{a_p\}, \{b_p\}$, we denote $a_p \lesssim b_p$ or $a_p = \mathcal{O}(b_p)$ if $\lim_{p \to \infty} a_p / b_p \leq c$, $a_p \gtrsim b_p$ or $a_p = \Omega(b_p)$ if $\lim_{p \to \infty} a_p / b_p \geq c$, for some constant $c > 0$, $a_p = o(b_p)$ if $\lim_{p \to \infty} a_p / b_p = 0$, and $a_p \asymp b_p$ if both $b_p \lesssim a_p$ and $a_p \lesssim b_p$. Throughout the paper, we use the terms "community" and "clusters" exchangeably.

3.2 Model formulation and motivations

Degree-corrected tensor block model

Suppose that we have an order-K data tensor $\mathcal{Y} \in \mathbb{R}^{p \times \cdots \times p}$. Assume that there exist $r \geq 1$ disjoint communities among the p nodes. We represent the community assignment by a function $z \colon [p] \mapsto [r]$, where $z(i) = a$ for i-th node that belongs to the a-th community. Then, $z^{-1}(a) = \{i \in [p] \colon z(i) = a\}$ denotes the set of nodes that belong to the a-th community, and $|z^{-1}(a)|$ denotes the number of nodes in the a-th community. Let $\boldsymbol{\theta} = (\theta(1), \ldots, \theta(p))^T$ denote the degree heterogeneity for p nodes. We consider the order-K dTBM (Ghoshdastidar and Dukkipati, 2017a; Ke et al., 2019),

$$\mathcal{Y}(i_1, \ldots, i_K) = \mathcal{S}(z(i_1), \ldots, z(i_K)) \prod_{k=1}^{K} \theta_{i_k} + \mathcal{E}(i_1, \ldots, i_K),$$

where $\mathcal{S} \in \mathbb{R}^{r\times\cdots\times r}$ is an order-K tensor collecting the block means among communities, and $\mathcal{E} \in \mathbb{R}^{p\times\cdots\times p}$ is a noise tensor consisting of independent zero-mean sub-Gaussian entries with variance bounded by σ^2. The unknown parameters are z, S, and $\boldsymbol{\theta}$. The dTBM can be equivalently written in a compact form of tensor-matrix product:

$$\mathbb{E}\mathcal{Y} = \mathcal{S} \times_1 \boldsymbol{\Theta}\boldsymbol{M} \times_2 \cdots \times_K \boldsymbol{\Theta}\boldsymbol{M}, \tag{3.1}$$

where $\boldsymbol{\Theta} = \text{diag}(\theta(1), ..., \theta(p)) \in \mathbb{R}^{p\times p}$ is a diagonal matrix, $\boldsymbol{M} \in \{0,1\}^{p\times r}$ is the membership matrix associated with community assignment z such that $\boldsymbol{M}(i,j) = \mathbb{1}\{z(i) = j\}$. By definition, each row of $\boldsymbol{M}$ has one copy of 1's and 0's elsewhere. Note that the discrete nature of $\boldsymbol{M}$ renders our model (3.1) more challenging than Tucker decomposition. We call a tensor $\mathcal{Y}$ an r-block tensor with degree $\boldsymbol{\theta}$ if $\mathcal{Y}$ admits dTBM (3.1) and let $\mathcal{X} = \mathbb{E}\mathcal{Y}$ denote the mean tensor. The goal of clustering is to estimate z from a single noisy tensor $\mathcal{Y}$. We are particularly interested in the high-dimensional regime where p grows whereas $r = \mathcal{O}(1)$.

For ease of notation, we have focused on the case with symmetric mean tensor $\mathbb{E}\mathcal{Y}$. This assumption simplifies the notation because all modes have the same $(\boldsymbol{\Theta}, \boldsymbol{M}, z)$; the noise tensor $\mathcal{E}$ and the data tensor $\mathcal{Y}$ are still possibly asymmetric. In general, we allow asymmetric mean tensors with $\{(\boldsymbol{\Theta}_k, \boldsymbol{M}_k, z_k)\}_{k=1}^K$, one for each mode. The extension can be found in Appendix B.

Motivating examples

Here, we provide four applications to illustrate the practical necessity of dTBM.

Tensor block model Consider the model (3.1). Let $\theta(i) = 1$ for all $i \in [p]$. The model (3.1) reduces to the tensor block model, which is widely used in previous clustering algorithms (Wang and Zeng, 2019; Chi et al., 2020; Han et al., 2022a). The theoretical results in TBM serve as benchmarks for dTBM.

Community detection in hypergraphs The hypergraph network is a powerful tool to represent the complex entity relations with higher-order interactions (Ke et al.,

2019). A typical undirected hypergraph is denoted as $H = (V, E)$, where $V = [p]$ is the set of nodes and E is the set of undirected hyperedges. Each hyperedge in E is a subset of V, and we call the hyperedge an order-K edge if the corresponding subset involves K nodes. We call H a K-uniform hypergraph if E only contains order-K edges.

It is natural to represent the K-uniform hypergraph using a binary order-K adjacency tensor. Let $\mathcal{Y} \in \{0,1\}^{p\times\cdots\times p}$ denote the adjacency tensor, where the entries encode the presence or absence of order-K edges among p nodes. Specifically, for all $(i_1, \ldots, i_K) \in [p]^K$, we have

$$\mathcal{Y}(i_1, ..., i_K) = \begin{cases} 1 & \text{if } (i_1, ..., i_K) \in E, \\ 0 & \text{if } (i_1, ..., i_K) \notin E. \end{cases}$$

Assume that there exist r disjoint communities among p nodes, and the connection probabilities depend on the community assignments and node-specific parameters. Then, the equation (3.1) models $\mathbb{E}\mathcal{Y}$ with unknown degree heterogeneity $\boldsymbol{\theta}$ and sub-Gaussianity parameter $\sigma^2 = 1/4$.

Multi-layer weighted network Multi-layer weighted network data consists of multiple networks over the same set of nodes. One representative example is the brain connectome data (Zhang et al., 2019). The multi-layer weighted network $\mathcal{Y}$ has dimension of $p \times p \times L$, where p denotes the number of brain regions of interest, and L denotes the number of layers (networks). Each of the L networks describes one aspect of the brain connectivity, such as functional connectivity or structural connectivity. The resulting tensor $\mathcal{Y}$ consists of a mixture of slices with various data types.

Assume that there exist r disjoint communities among p nodes and r_l disjoint communities among the L layers. The multi-layer network community detection is modeled by the general asymmetric dTBM model (3.1)

$$\mathbb{E}\mathcal{Y} = \mathcal{S} \times_1 \boldsymbol{\Theta}\boldsymbol{M} \times_2 \boldsymbol{\Theta}\boldsymbol{M} \times_3 \boldsymbol{\Theta}_l\boldsymbol{M}_l,$$

where ($\boldsymbol{\theta} \in \mathbb{R}^p, \boldsymbol{M} \in \{0,1\}^{p\times r}$) and ($\boldsymbol{\theta}_l \in \mathbb{R}^L, \boldsymbol{M}_l \in \{0,1\}^{L\times r_l}$) are the degree heterogeneity and membership matrices corresponding to the community structure for p nodes and L layers, respectively.

Gaussian higher-order clustering Datasets in various fields such as medical image, genetics, and computer science are formulated as Gaussian tensors. One typical example is the multi-tissue gene expression dataset, which records different gene expressions in different individuals and different tissues. The dataset, denoted as $\mathcal{Y} \in \mathbb{R}^{p\times n\times t}$, consists of the expression data for p genes of n individuals in t tissues.

Assume that there exist r_1, r_2, r_3 disjoint clusters for p genes, n individuals, and t tissues, respectively. We apply the general asymmetric dTBM model (3.1)

$$\mathbb{E}\mathcal{Y} = \mathcal{S} \times_1 \boldsymbol{\Theta}_1\boldsymbol{M}_1 \times_2 \boldsymbol{\Theta}_2\boldsymbol{M}_2 \times_3 \boldsymbol{\Theta}_3\boldsymbol{M}_3,$$

where $\{(\boldsymbol{\theta}_k, \boldsymbol{M}_k)\}_{k=1}^3$ represents the degree heterogeneity and membership for genes, individuals, and tissues.

Remark 3.1 (Comparison with non-degree models). Our dTBM uses fewer block parameters than TBM. In particular, every non-degree r_1-block tensor can be represented by a *degree-corrected* r_2-block tensor with $r_2 \le r_1$. In particular, there exist tensors with $r_1 = p$ but $r_2 = 1$, so the reduction in model complexity can be dramatic from p to 1. This fact highlights the benefits of introducing degree heterogeneity in higher-order clustering tasks.

Identifiability under angle gap condition

The goal of clustering is to estimate the partition function z from model (3.1). For ease of notation, we focus on symmetric tensors; the extension to non-symmetric tensors are similar. We use $\mathcal{P}$ to denote the following parameter space for $(z, \mathcal{S}, \boldsymbol{\theta})$,

$$\mathcal{P} = \Big\{ (z, \mathcal{S}, \boldsymbol{\theta}) \colon \ \boldsymbol{\theta} \in \mathbb{R}_+^p, \ \frac{c_1 p}{r} \le |z^{-1}(a)| \le \frac{c_2 p}{r}, \\ c_3 \le \|\mathrm{Mat}(\mathcal{S})_{a:}\| \le c_4, \left\|\boldsymbol{\theta}_{z^{-1}(a)}\right\|_1 = |z^{-1}(a)|, a \in [r] \Big\} \tag{3.2}$$

where $c_i > 0$'s are universal constants. We briefly describe the rationale of the constraints in (3.2). First, the entrywise positivity constraint on $\boldsymbol{\theta} \in \mathbb{R}_+^p$ is imposed to avoid sign ambiguity between entries in $\boldsymbol{\theta}_{z^{-1}(a)}$ and $\mathcal{S}$. This constraint allows the trigonometric $\cos$ to describe the angle similarity in the Assumption 3.2 below and Sub-algorithm 2 in Section 3.4. Note that the positivity constraint can be achieved without sacrificing model flexibility, by using a slightly larger dimension of $\mathcal{S}$ in the factorization (3.1); see Example 3.1 below. Second, recall that the quantity $|z^{-1}(a)|$ denotes the number of nodes in the a-th community. The constants c_1, c_2 in the $|z^{-1}(a)|$ bounds assume the roughly balanced size across r communities. Third, the constant c_3 requires that all slides in $\mathcal{S}$ have non-degenerate norm. Particularly, the lower bound c_3 excludes the purely zero slide to avoid trivial non-identifiability of model (3.1); see Example 3.2 below. The upper bound c_4 is a technical constraint to avoid the slides with diverging norm as dimension grows. Lastly, the ℓ_1 normalization $\|\boldsymbol{\theta}_{z^{-1}(a)}\|_1 = |z^{-1}(a)|$ is imposed to avoid the scalar ambiguity between $\boldsymbol{\theta}_{z^{-1}(a)}$ and $\mathcal{S}$. This constraint, again, incurs no restriction to model flexibility but makes our presentation cleaner. Our constraints in $\mathcal{P}$ are mild compared with previous literature; see Table 3.2 for comparison.

Example 3.1 (Positivity of degree parameters)**.** *Here we provide an example to show the positivity constraint on $\boldsymbol{\theta}$ incurs no loss on the model flexibility. Consider an order-3 dTBM with core tensor $\mathcal{S} = 1$ and degree $\boldsymbol{\theta} = (1, 1, -1, -1)^T$. We have the mean tensor*

$$\mathcal{X} = \mathcal{S} \times_1 \boldsymbol{\Theta}\boldsymbol{M} \times_2 \boldsymbol{\Theta}\boldsymbol{M} \times_3 \boldsymbol{\Theta}\boldsymbol{M},$$

where $\boldsymbol{\Theta} = diag(\boldsymbol{\theta})$ and $\boldsymbol{M} = (1, 1, 1, 1)^T$. Note that $\mathcal{X} \in \mathbb{R}^{4\times4\times4}$ is a 1-block tensor with

mixed-signed *degree* $\boldsymbol{\theta}$*, and the mode-3 slices of* $\mathcal{X}$ *are*

$$\mathcal{X}_{::1} = \mathcal{X}_{::2} = -\mathcal{X}_{::3} = -\mathcal{X}_{::4} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}.$$

Now, instead of original decomposition, we encode $\mathcal{X}$ *as a 2-block tensor with* positive-signed *degree. Specifically, we write*

$$\mathcal{X} = \mathcal{S}' \times_1 \boldsymbol{\Theta}'\boldsymbol{M}' \times_2 \boldsymbol{\Theta}'\boldsymbol{M}' \times_3 \boldsymbol{\Theta}'\boldsymbol{M}',$$

where $\boldsymbol{\Theta}' = \mathit{diag}(\boldsymbol{\theta}') = \mathit{diag}(1,1,1,1)$*, the core tensor* $\mathcal{S}' \in \mathbb{R}^{2\times 2\times 2}$ *has following mode-3 slices, and the membership matrix* $\boldsymbol{M}' \in \{0,1\}^{4\times 2}$ *defines the clustering* $z'\colon [4] \to [2]$*; i.e.,*

$$\mathcal{S}'_{::1} = -\mathcal{S}'_{::2} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad \boldsymbol{M}' = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

The triplet $(z', \mathcal{S}', \boldsymbol{\theta}')$ *lies in our parameter space* (3.2)*. In general, we can always reparameterize an* r*-block tensor with mixed-signed degree using a* $2r$*-block tensor with positive-signed degree. Since we assume* $r = \mathcal{O}(1)$ *throughout the paper, the splitting does not affect the error rates of our interest.*

Example 3.2 (Non-identifiability with purely zero core slice)**.** *Consider an order-2 dTBM with core tensor* $\boldsymbol{S} = \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix}$ *degree matrices* $\boldsymbol{\Theta}_1 = \boldsymbol{\Theta}_2 = \mathit{diag}(1,1,1,1)$*, and mean tensor*

$$\mathcal{X} = \boldsymbol{\Theta}_1\boldsymbol{M}\boldsymbol{S}\boldsymbol{M}^T\boldsymbol{\Theta}_2, \quad \textit{with } \boldsymbol{M} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

Replacing $\mathbf{\Theta}_1$ *by* $\mathbf{\Theta}_1' = (3/2, 1/2, 1, 1)$ *leads to the same mean tensor* $\mathcal{X}$.

Assumptions in parameter space	Gao et al. (2018)	Han et al. (2022a)	Ke et al. (2019)	**Ours**
Balanced community sizes	√	√	√	√
Bounded core tensors	√	×	√	√
Balanced degrees	√	-	√	√
Flexible in-group connections	×	√	√	√
Gaps among cluster centers	In-between cluster difference	Euclidean gap	Eigen gap	Angle gap

Table 3.2: Parameter space comparison between previous work with our assumption.

We now provide the identifiability conditions for our model before estimation procedures. When $r = 1$, the decomposition (3.1) is always unique (up to cluster label permutation) in $\mathcal{P}$, because dTBM is equivalent to the rank-1 tensor family under this case. When $r \geq 2$, the Tucker rank of signal tensor $\mathbb{E}\mathcal{Y}$ in (3.1) is bounded by, but not necessarily equal to, the number of blocks r (Wang and Zeng, 2019). Therefore, one can not apply the classical identifiability conditions for low-rank tensors to dTBM. Here, we introduce a key separation condition on the core tensor.

Assumption 3.2 (Angle gap)**.** *Let* $\boldsymbol{S} = Mat(\mathcal{S})$*. Assume that the minimal gap between normalized rows of* $\boldsymbol{S}$ *is bounded away from zero; i.e.,*

$$\Delta_{\min} := \min_{a \neq b \in [r]} \left\| \frac{\boldsymbol{S}_{a:}}{\|\boldsymbol{S}_{a:}\|} - \frac{\boldsymbol{S}_{b:}}{\|\boldsymbol{S}_{b:}\|} \right\| > 0, \quad \textit{for} \quad r \geq 2. \tag{3.3}$$

We make the convention $\Delta_{\min} = 1$ for $r = 1$. Equivalently, (3.3) says that none of the two rows in $\boldsymbol{S}$ are parallel; i.e., $\max_{a \neq b \in [r]} \cos(\boldsymbol{S}_{a:}, \boldsymbol{S}_{b:}) = 1 - \Delta_{\min}^2/2 < 1$. The quantity $\Delta_{\min}$ characterizes the non-redundancy among clusters measured by angle separation. The denominators involved in definition (3.3) are well posed because of the lower bound on $\|\boldsymbol{S}_{a:}\|$ in (3.2).

Our first main result is the following theorem showing the sufficiency and necessity of the angle gap separation condition for the parameter identifiability under dTBM.

Theorem 3.3 (Model identifiability)**.** *Consider the dTBM with $r \geq 2$ and $K \geq 2$. The parameterization* (3.1) *is unique in $\mathcal{P}$ up to cluster label permutations, if and only if Assumption 3.2 holds.*

The identifiability guarantee for the dTBM is stronger than classical Tucker model. In the Tucker model, the factor matrix $\boldsymbol{M}$ is identifiable only up to orthogonal rotations. In contrast, our model does not suffer from rotational invariance. As we will show in Section 3.4, each column of the membership matrix $\boldsymbol{M}$ can be precisely recovered under our algorithm. This property benefits the interpretation of dTBM in practice.

3.3 Statistical-computational critical values for higher-order tensors

Assumptions

We propose the signal-to-noise ratio (SNR),

$$\text{SNR} := \Delta_{\min}^2/\sigma^2 = p^\gamma, \tag{3.4}$$

with varying $\gamma \in \mathbb{R}$ that quantifies different regimes of interest. We call γ the *signal exponent*. Intuitively, a larger SNR, or equivalently a larger γ, benefits the clustering in the presence of noise. With quantification (3.4), we consider the following parameter space,

$$\mathcal{P}(\gamma) = \mathcal{P} \cap \{\mathcal{S} \text{ satisfies SNR condition (3.4) with } \gamma\}. \tag{3.5}$$

The 1-block dTBM does not belong to the space $\mathcal{P}(\gamma)$ when $\gamma < 0$, due to the convention in Assumption 3.2. Our goal is to characterize the clustering accuracy with respect to γ under the space $\mathcal{P}(\gamma)$.

In our algorithmic development, we often refer to the regime of balanced degree

heterogeneity. We call the degree $\boldsymbol{\theta}$ *balanced* if

$$\min_{a\in[r]}\|\boldsymbol{\theta}_{z^{-1}(a)}\| = (1+o(1))\max_{a\in[r]}\|\boldsymbol{\theta}_{z^{-1}(a)}\|. \tag{3.6}$$

The following lemma provides the rationale of balanced degree assumption. We show the close relation between angle gaps in the mean tensor $\mathcal{X}$ and the core tensor $\mathcal{S}$ under balanced degree heterogeneity.

Lemma 3.4 (Angle gaps in $\mathcal{X}$ and $\mathcal{S}$)**.** *Consider the dTBM model* (3.1) *under the parameter space* $\mathcal{P}$ *in* (3.2) *with* $r \geq 2$*. Suppose* $\boldsymbol{\theta}$ *is balanced satisfying* (3.6) *and* $\min_{i\in[p]}\theta(i) \geq c$ *from some constant* $c > 0$*. Then, as* $p \to \infty$*, for all* i, j *such that* $z(i) \neq z(j)$*, we have*

$$\cos(\boldsymbol{X}_{i:}, \boldsymbol{X}_{j:}) \asymp \cos(\boldsymbol{S}_{z(i):}, \boldsymbol{S}_{z(j):}),$$

where $\boldsymbol{X} = Mat(\mathcal{X})$ *and* $\boldsymbol{S} = Mat(\mathcal{S})$*.*

In practice, an estimation algorithm has access to a noisy version of $\mathcal{X}$ but not $\mathcal{S}$. Our goal is to establish the algorithm performance with respect to the signal $\Delta^2_{\min}$ in the core tensor. By Lemma 3.4, the mapping from the core tensor $\boldsymbol{S}_{z(i):}$ to the mean tensor $\boldsymbol{X}_{z(i):}$ preserves the angle information $\Delta^2_{\min}$ under balanced degree heterogeneity (3.6). Therefore, the balanced degree assumption helps to exclude the cases in which the degree heterogeneity distorts the algorithm guarantees.

Here, we provide an example to illustrate the insufficiency of $\Delta^2_{\min}$ in the absence of balanced degrees.

Example 3.3 (Insufficiency of $\Delta^2_{\min}$ in the absence of balanced degrees)**.** *Consider an order-2* (p, p)*-dimensional dTBM with core matrix*

$$\boldsymbol{S} = \begin{pmatrix} 1 & a \\ 1 & -a \end{pmatrix}, \tag{3.7}$$

and $\boldsymbol{\theta}$ *such that* $\|\boldsymbol{\theta}_{z^{-1}(1)}\|^2 = p^m\|\boldsymbol{\theta}_{z^{-1}(2)}\|^2$*, where* $m \in [-1, 1]$ *is a scalar parameter controlling the skewness of degrees. Let* $\Delta^2_{\boldsymbol{X}}$ *denote the minimal angle gap of the mean*

tensor, defined by

$$\Delta_{\boldsymbol{X}}^2 \coloneqq \min_{i,j\in[p],z(i)\neq z(j)} \left\| \frac{\boldsymbol{X}_{i:}}{\|\boldsymbol{X}_{i:}\|} - \frac{\boldsymbol{X}_{j:}}{\|\boldsymbol{X}_{j:}\|} \right\|, \tag{3.8}$$

where $\boldsymbol{X} = Mat(\mathcal{X})$. Take $a = p^{-1/4}$ in the model setup (3.7)*. We have*

$$\Delta_{\min}^2 = \frac{2a^2}{1+a^2} \asymp p^{-1/2},$$
$$\Delta_{\boldsymbol{X}}^2 = \frac{2\|\boldsymbol{\theta}_{z^{-1}(2)}\|^2 a^2}{\|\boldsymbol{\theta}_{z^{-1}(1)}\|^2 + \|\boldsymbol{\theta}_{z^{-1}(2)}\|^2 a^2} \asymp p^{-1/2-m}.$$

Based on the Theorem 3.7 in Section 3.3, the dTBM is impossible to solve when $\Delta_{\boldsymbol{X}}^2 \lesssim p^{-1}$ even though $\Delta_{\min}^2 \asymp p^{-1/2}$; that is, the dTBM estimation depends on the relative magnitude of m vs. $1/2$. In such a setting, the proposed signal notion $\Delta_{\min}^2$ alone fails to fully characterize dTBM.

Remark 3.5 (Flexibility in balanced degree assumption). One important note is that our balance assumption (3.6) does not preclude the mild degree heterogeneity. In fact, within each of the clusters, we allow the highest degree at the order $\mathcal{O}(p)$, whereas the lowest degree at the order $\Omega(1)$. This range is more relaxed than previous work (Gao et al., 2018) that restricts the highest degree in the sub-linear regime $o(p)$ and the lowest degree at the order $\Omega(1)$.

Remark 3.6 (Similar assumptions in literature). Similar degree regulations are not rare in literature. In higher-order tensor model (Ke et al., 2019), the degree assumption $\max_{a\in[r]}\|\boldsymbol{\theta}_{z^{-1}(a)}\| \leq C \min_{a\in[r]}\|\boldsymbol{\theta}_{z^{-1}(a)}\|$ is made to ensure degree balance across communities. In Gao et al. (2018), the degree distribution is restricted to $\frac{1}{|z^{-1}(a)|}\sum_{i\in z^{-1}(a)}\theta_i = 1 + o(1)$ for all communities.

Last, let $\hat{z}$ and z be the estimated and true clustering functions in the family (3.2). Define the misclustering error by

$$\ell(\hat{z}, z) = \frac{1}{p}\min_{\pi\in\Pi}\sum_{i\in[p]} \mathbb{1}\{\hat{z}(i) \neq \pi\circ z(i)\},$$

where $\pi : [r] \mapsto [r]$ is a permutation of cluster labels, $\circ$ denotes the composition

operation, and Π denotes the collection of all possible permutations. The infimum over all permutations accounts for the ambiguity in cluster label permutation.

In Sections 3.3 and 3.3, we provide the phase transition of $\ell(\hat{z}, z)$ for general Gaussian dTBMs (3.1) without symmetric assumptions. For general (asymmetric) Gaussian dTBMs, we assume Gaussian noise $\mathcal{E}(i_1, \ldots, i_K) \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$, and we extend the parameter space (3.2) to allow K clustering functions $\{z_k\}_{k\in[K]}$, one for each mode. For notational simplicity, we still use z and $\mathcal{P}(\gamma)$ for this general (asymmetric) model. All results should be interpreted as the worst-case results across K modes.

Statistical critical value

The statistical critical value means the SNR required for solving dTBMs with *unlimited computational cost*. Our following result shows the minimax lower bound for exact recovery and the matching upper bound for maximum likelihood estimator (MLE). We consider the Gaussian MLE, denoted as $(\hat{z}_{\text{MLE}}, \hat{\mathcal{S}}_{\text{MLE}}, \hat{\boldsymbol{\theta}}_{\text{MLE}})$, over the estimation space $\mathcal{P}$, where

$$(\hat{z}_{\text{MLE}}, \hat{\mathcal{S}}_{\text{MLE}}, \hat{\boldsymbol{\theta}}_{\text{MLE}}) = \underset{(z,\mathcal{S},\boldsymbol{\theta})\in\mathcal{P}}{\arg\min} \|\mathcal{Y} - \mathcal{X}(z, \mathcal{S}, \boldsymbol{\theta})\|_F^2. \tag{3.9}$$

Theorem 3.7 (Statistical critical value)**.** *Consider general Gaussian dTBMs with parameter space $\mathcal{P}(\gamma)$ and $K \geq 2$. Then, we have the following statistical phase transition.*

- ***Impossibility.*** *Assume $p \to \infty$ and $2 \leq r \lesssim p^{1/3}$. Let $\mathcal{P}_{\mathcal{S}}(\gamma) \coloneqq \{\mathcal{S} : c_3 \leq \|Mat(\mathcal{S})_{a:}\| \leq c_4, a \in [r]\} \cap \{\mathcal{S} : \Delta_{\min}^2 = p^{\gamma}\}$ denote the space for valid $\mathcal{S}$ satisfying SNR condition* (3.4)*, and $\mathcal{P}_{z,\boldsymbol{\theta}} \coloneqq \{\boldsymbol{\theta} \in \mathbb{R}_+^p,\ \frac{c_1 p}{r} \leq |z^{-1}(a)| \leq \frac{c_2 p}{r}, \left\|\boldsymbol{\theta}_{z^{-1}(a)}\right\|_1 = |z^{-1}(a)|, a \in [r]\}$ denote the space for valid $(z, \boldsymbol{\theta})$, where c_1, c_2, c_3, c_4 are the constants in parameter space* (3.2)*. If the signal exponent satisfies $\gamma < -(K-1)$, then, for any true core tensor $\mathcal{S} \in \mathcal{P}_{\mathcal{S}}(\gamma)$, no estimator $\hat{z}_{stat}$ achieves exact recovery in expectation; that is, when $\gamma < -(K-1)$, we have*

$$\liminf_{p\to\infty} \inf_{\mathcal{S}\in\mathcal{P}_{\mathcal{S}}(\gamma)} \inf_{\hat{z}_{stat}} \sup_{(z,\boldsymbol{\theta})\in\mathcal{P}_{z,\boldsymbol{\theta}}} \mathbb{E}\left[p\ell(\hat{z}_{stat}, z)\right] \geq 1. \tag{3.10}$$

Further, we define the parameter space $\mathcal{P}'(\gamma') \coloneqq \mathcal{P} \cap \{\Delta_{\boldsymbol{X}}^2 = p^{\gamma'}\}$, *where* $\Delta_{\boldsymbol{X}}^2$ *is the mean tensor minimal gap in* (3.8). *When* $\gamma' < -(K-1)$, *we have*

$$\liminf_{p\to\infty} \inf_{\hat{z}_{stat}} \sup_{(z,\mathcal{S},\boldsymbol{\theta})\in\mathcal{P}'(\gamma')} \mathbb{E}\left[p\ell(\hat{z}_{stat}, z)\right] \geq 1.$$

- ***MLE achievability.*** *Suppose that the signal exponent satisfies* $\gamma > -(K-1) + c_0$ *for an arbitrary constant* $c_0 > 0$. *Furthermore, assume that* $\boldsymbol{\theta}$ *is balanced and* $\min_{i\in[p]} \theta(i) \geq c$ *from some constant* $c > 0$. *Then, when* $p \to \infty$, *for fixed* $r \geq 1$, *the MLE in* (3.9) *achieves exact recovery in high probability; that is,*

$$\ell(\hat{z}_{MLE}, z) \lesssim SNR^{-1} \exp\left(-\frac{p^{K-1}SNR}{r^{K-1}}\right) \to 0,$$

with probability going to 1.

The proofs for the two parts in Theorem 3.7 are in the Appendix B, Section B.2 and Section B.2, respectively. The first part of Theorem 3.7 demonstrates impossibility of exact recovery whenever the core tensor $\mathcal{S}$ satisfies SNR condition (3.4) with exponent $\gamma < -(K-1)$. The proof is information-theoretical, and therefore the results apply to all statistical estimators, including but not limited to MLE and trace maximization (Ghoshdastidar and Dukkipati, 2017b). The minimax bound (3.10) indicates the worst case impossibility for a particular core tensor $\mathcal{S}$ with signal exponent $\gamma < -(K-1)$; i.e., under the assumptions of Theorem 3.7, when $\gamma < -(K-1)$, we have

$$\liminf_{p\to\infty} \inf_{\hat{z}_{\text{stat}}} \sup_{(z,\mathcal{S},\boldsymbol{\theta})\in\mathcal{P}(\gamma)} \mathbb{E}\left[p\ell(\hat{z}_{\text{stat}}, z)\right] \geq 1.$$

Such worst case impossibility is studied in related works (Han et al., 2022a; Gao et al., 2018) while our lower bound (3.10) provides a stronger impossibility statement for arbitrary core tensors with weak signals. The second part of Theorem 3.7 shows the exact recovery of MLE when $\gamma > -(K-1) + c_0$ for an arbitrary constant $c_0 > 0$. Combining the impossibility and achievability results, we conclude that the boundary $\gamma_{\text{stat}} \coloneqq -(K-1)$ is the critical value for statistical performance of dTBM

with respect to our SNR.

Computational critical value

The computational critical value means the minimal SNR required for exact recovery with *polynomial-time* computational cost. An important ingredient to establish the computational limits is the *hypergraphic planted clique* (*HPC*) *conjecture* (Zhang and Xia, 2018; Brennan and Bresler, 2020). The HPC conjecture indicates the impossibility of fully recovering the planted cliques with polynomial-time algorithm when the clique size is less than the number of vertices in the hypergraph. The formal statement of HPC detection conjecture is provided in Definition 3.8 and Conjecture 3.4 as follows.

Definition 3.8 (Hypergraphic planted clique (HPC) detection)**.** *Consider an order-K hypergraph $H = (V, E)$ where $V = [p]$ collects vertices and E collects all the order-K edges. Let $\mathcal{H}_k(p, 1/2)$ denote the Erdős-Rényi K-hypergraph where the edge $(i_1, \ldots, i_K)$ belongs to E with probability $1/2$. Further, we let $\mathcal{H}_K(p, 1/2, \kappa)$ denote the hyhpergraph with planted cliques of size κ. Specifically, we generate a hypergraph from $\mathcal{H}_k(p, 1/2)$, pick κ vertices uniformly from $[p]$, denoted K, and then connect all the hyperedges with vertices in K. Note that the clique size κ can be a function of p, denoted κ_p. The order-K HPC detection aims to identify whether there exists a planted clique hidden in an Erdős-Rényi K-hypergraph. The HPC detection is formulated as the following hypothesis testing problem*

$$H_0 : \ H \sim \mathcal{H}_K(p, 1/2) \quad \textit{versus} \quad H_1 : \ H \sim \mathcal{H}_K(p, 1/2, \kappa_p).$$

Conjecture 3.4 (HPC conjecture)**.** *Consider the HPC detection problem in Definition 3.8 with $K \geq 2$. Suppose the sequence $\{\kappa_p\}$ such that $\limsup_{p\to\infty} \log \kappa_p / \log \sqrt{p} \leq (1 - \tau)$ for any $\tau > 0$. Then, for every sequence of polynomial-time test $\{\varphi_p\} : H \mapsto \{0, 1\}$ we have*

$$\liminf_{p\to\infty} \mathbb{P}_{H_0} (\varphi_p(H) = 1) + \mathbb{P}_{H_1} (\varphi_p(H) = 0) > \frac{1}{2}.$$

Under the HPC conjecture, we establish the SNR lower bound that is necessary for any *polynomial-time* estimator to achieve exact clustering.

Theorem 3.9 (Computational critical value)**.** *Consider general Gaussian dTBMs under the parameter space $\mathcal{P}$ with $K \geq 2$. Then, we have the following computational phase transition.*

- ***Impossibility.*** *Assume HPC conjecture holds and $r \geq 2$. If the signal exponent satisfies $\gamma < -K/2$, then, no* polynomial-time estimator *$\hat{z}_{comp}$ achieves exact recovery in expectation as $p \to \infty$; that is, when $\gamma < -K/2$, we have*

$$\liminf_{p\to\infty} \sup_{(z,\mathcal{S},\boldsymbol{\theta})\in\mathcal{P}(\gamma)} \mathbb{E}\left[p\ell(\hat{z}_{comp}, z)\right] \geq 1.$$

- ***Polynomial-time algorithm achievability.*** *Suppose that we have fixed $r \geq 1$, and the signal exponent satisfies $\gamma > -K/2 + c_0$ for an arbitrary constant $c_0 > 0$. Furthermore, assume that the degree $\boldsymbol{\theta}$ is balanced, lower bounded in that $\min_{i\in[p]} \theta_i \geq c$ for some constant $c > 0$, and satisfies the locally linear stability in Definition 3.16 in the neighborhood $\mathcal{N}(z, \varepsilon)$ for all $\varepsilon \leq E_0$ and some $E_0 \gtrsim \log^{-1} p$. Then, as $p \to \infty$, there exists a polynomial-time algorithm $\hat{z}_{ploy}$ that achieves exact recovery in high probability; that is,*

$$\ell(\hat{z}_{poly}, z) \lesssim SNR^{-1} \exp\left(-\frac{p^{K-1}SNR}{r^{K-1}}\right) \to 0,$$

with probability going to 1.

The proofs for the two parts in Theorem 3.9 are in the Appendix B, Section B.2 and Section B.2, respectively. The first part of Theorem 3.9 indicates the impossibility of exact recovery by polynomial-time algorithms when $\gamma < -K/2$, and the second part shows the existence of such algorithm when $\gamma > -K/2 + c_0$ for an arbitrary constant $c_0 > 0$ under extra technical assumptions. In Section 3.4, we will present an efficient polynomial-time algorithm in this setting. Therefore, we conclude that $\gamma_{\text{comp}} := -K/2$ is the critical value for computational performance of dTBM with respect to our SNR.

Remark 3.10 (Statistical-computational gaps). Now, we have established the phase transition of exact clustering under order-K dTBM by combining Theorems 3.7 and 3.9. Figure 3.2 summarizes our results of critical SNRs when $K \geq 2$. In the weak SNR region $\gamma < -(K-1)$, no statistical estimator succeeds in degree-corrected higher-order clustering. In the strong SNR region $\gamma > -K/2$, our proposed algorithm precisely recovers the clustering in polynomial time. In the moderate SNR regime, $-(K-1) \leq \gamma \leq -K/2$, the degree-corrected clustering problem is statistically easy but computationally hard. Particularly, dTBM reduces to matrix degree-corrected model when $K = 2$, and the statistical and computational bounds show the same critical value. When $K = 1$, dTBM reduces to the degree-corrected sub-Gaussian mixture model (GMM) with model

$$\boldsymbol{Y} = \boldsymbol{\Theta}\boldsymbol{M}\boldsymbol{S} + \boldsymbol{E},$$

where $\boldsymbol{Y} \in \mathbb{R}^{p\times d}$ collects n data points in $\mathbb{R}^d$, $\boldsymbol{S} \in \mathbb{R}^{r\times d}$ collects the d-dimensional centroids for r clusters, and $\boldsymbol{\Theta} \in \mathbb{R}^{p\times p}, \boldsymbol{M} \in \{0,1\}^{p\times r}, \boldsymbol{E} \in \mathbb{R}^{p\times d}$ have the same meaning as in dTBM. Lu and Zhou (2016) implies that polynomial-time algorithms are able to achieve the statistical minimax lower bound in GMM. Therefore, we conclude that the statistical-computational gap emerges only for higher-order tensors with $K \geq 3$. The result reveals the intrinsic distinctions among (vector) one-dimensional clustering, (matrix) biclustering, and (tensor) higher-order clustering.

3.4 Polynomial-time algorithm under mild SNR

In this section, we present an efficient polynomial-time clustering algorithm under mild SNR. The procedure takes a global-to-local approach. See Figure 3.3 for illustration. The global step finds the basin of attraction with polynomial misclustering error, whereas the local iterations improve the initial clustering to exact recovery. Both steps are critical to obtain a satisfactory algorithm output. In what follows, we first use the symmetric tensor as a working example to describe the algorithm pro-

cedures to gain insight. Our theoretical analysis focuses on dTBMs with symmetric mean tensor and independent sub-Gaussian noises such as Gaussian and uniform observations. The extensions for Bernoulli observations and other practical issues are in Sections 3.4 and 3.4.

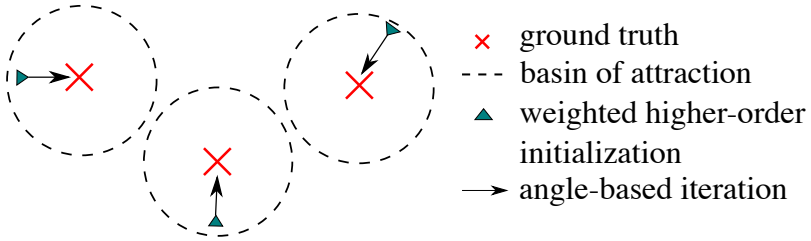

Figure 3.3: Illustration of global-to-local dTBM algorithm.

Illustration of our global-to-local algorithm.

To construct algorithm guarantees, we introduce the misclustering loss between an estimator $\hat{z}$ and the true z:

$$L(\hat{z}, z) = \frac{1}{p} \sum_{i \in [p]} \theta(i) \sum_{b \in [r]} \mathbb{1}\{\hat{z}(i) = b\} \cdot \left\| \left[\boldsymbol{S}_{z(i):} \right]^s - [\boldsymbol{S}_{b:}]^s \right\|^2, \tag{3.11}$$

where the superscript $\cdot^s$ denotes the normalized vector; i.e., $\boldsymbol{a}^s := \boldsymbol{a}/\|\boldsymbol{a}\|$ if $\boldsymbol{a} \neq 0$ and $\boldsymbol{a}^s = 0$ if $\boldsymbol{a} = 0$ for any vector $\boldsymbol{a}$. The following lemma indicates the close relationship between the loss $L(\hat{z}, z)$ and error $\ell(\hat{z}, z)$. The loss $L(\hat{z}, z)$ serves as an important intermediate quantity to control the misclustering error.

Lemma 3.11 (Relationship between misclustering error and loss)**.** *Consider the dTBM under the parameter space $\mathcal{P}$. Suppose $\min_{i \in [p]} \theta(i) > c$ for some constant $c > 0$. We have $\ell(\hat{z}, z)\Delta_{\min}^2 \leq L(\hat{z}, z)$.*

Weighted higher-order initialization

We start with weighted higher-order clustering algorithm as initialization. We take an order-3 tensor and the clustering on the first mode as illustration for insight.

Consider noiseless case with $\mathcal{X} = \mathbb{E}\mathcal{Y}$ and $\boldsymbol{X} = \text{Mat}(\mathcal{X})$. By model (3.1), for all $i \in [p]$, we have

$$\theta(i)^{-1}\boldsymbol{X}_{i:} = [\text{Mat}(\mathcal{S} \times_2 \boldsymbol{\Theta}\boldsymbol{M} \times_3 \boldsymbol{\Theta}\boldsymbol{M})]_{z(i):}.$$

This implies that, all node i belonging to the a-th community (i.e., $z(i) = a$) share the same normalized mean vector $\theta(i)^{-1}\boldsymbol{X}_{i:}$, and vice versa. Intuitively, one can apply k-means clustering to the vectors $\{\theta(i)^{-1}\boldsymbol{X}_{i:}\}_{i\in[p]}$, which leads to main idea of our Sub-algorithm 1.

Specifically, our initialization consists of the denoising step and the clustering step. The denoising step (lines 1-2 in Sub-algorithm 1) estimates $\mathcal{X}$ from $\mathcal{Y}$ by a double projection spectral method. The first projection performs HOSVD (De Lathauwer et al., 2000) via $\boldsymbol{U}_{\text{pre},k} = \text{SVD}_r\left(\text{Mat}_k(\mathcal{Y})\right), k \in [3]$, where $\text{SVD}_r(\cdot)$ returns the top-r left singular vectors. The second projection performs HOSVD on the projected $\mathcal{Y}$ onto the multilinear Kronecker space $\boldsymbol{U}_{\text{pre},k} \otimes \boldsymbol{U}_{\text{pre},k}$; i.e.,

$$\hat{\boldsymbol{U}}_1 = \text{SVD}_r\left(\text{Mat}_1\left(\mathcal{Y} \times_2 \boldsymbol{U}_{\text{pre},2}\boldsymbol{U}_{\text{pre},2}^T \times_3 \boldsymbol{U}_{\text{pre},3}\boldsymbol{U}_{\text{pre},3}^T\right)\right).$$

and similar for $\hat{\boldsymbol{U}}_2, \hat{\boldsymbol{U}}_3$. The final denoised tensor $\hat{\mathcal{X}}$ is defined by

$$\hat{\mathcal{X}} = \mathcal{Y} \times_1 \hat{\boldsymbol{U}}_1\hat{\boldsymbol{U}}_1^T \times_2 \hat{\boldsymbol{U}}_2\hat{\boldsymbol{U}}_2^T \times_3 \hat{\boldsymbol{U}}_3\hat{\boldsymbol{U}}_3^T.$$

The double projection improves usual matrix spectral methods in order to alleviate the noise effects for $K \geq 3$ (Han et al., 2022a).

The clustering step (lines 3-5 in Sub-algorithm 1) performs the weighted k-means clustering. We write $\hat{\boldsymbol{X}} = \text{Mat}_1(\hat{\mathcal{X}})$, and normalize the rows into $\hat{\boldsymbol{X}}_{i:}^s = \|\hat{\boldsymbol{X}}_{i:}\|^{-1}\hat{\boldsymbol{X}}_{i:}$ as a surrogate of $\theta(i)^{-1}\boldsymbol{X}_{i:}$. Then, a weighted k-means clustering is performed on the normalized rows with weights equal to $\|\hat{\boldsymbol{X}}_{i:}\|^2$. The choice of weights is to bound the k-means objective function by the Frobenius-norm accuracy of $\hat{\mathcal{X}}$. Unlike existing clustering algorithm (Ke et al., 2019), we apply the clustering on the unfolded tensor $\hat{\boldsymbol{X}}$ rather than on the factors $\hat{\boldsymbol{U}}_k$. This strategy relaxes the singular-value gap condition (Gao et al., 2018; Han et al., 2022a). We assign degenerate rows with purely zero entries to an arbitrarily random cluster; these

nodes are negligible in high-dimensions because of the lower bound on $\|\text{Mat}(\mathcal{S})_{a:}\|$ in (3.2). The final result gives the initial cluster assignment $z^{(0)}$. Full procedures for clustering are provided in Sub-algorithm 1.

We now establish the misclustering error rate of initialization.

Theorem 3.12 (Error for weighted higher-order initialization)**.** *Consider the general sub-Gaussian dTBM with fixed $r \geq 1$, $K \geq 2$, i.i.d. noise under the parameter space $\mathcal{P}$, and Assumption 3.2. Assume $\min_{i\in[p]} \theta(i) \geq c$ for some constant $c > 0$. Let $\Delta_{\boldsymbol{X}}$ denote the minimal gap in mean tensor defined in* (3.8)*, and let $z_k^{(0)}$ denote the output of Sub-algorithm 1. With probability going to 1, as $p \to \infty$, we have*

$$\ell(z_k^{(0)}, z) \lesssim \frac{\sigma^2 r^K p^{-K/2}}{\Delta_{\boldsymbol{X}}^2}.$$

Further, assume that $\boldsymbol{\theta}$ is balanced as (3.6)*. We have*

$$\ell(z_k^{(0)}, z) \lesssim \frac{r^K p^{-K/2}}{SNR} \text{ and } L(z_k^{(0)}, z) \lesssim \sigma^2 r^K p^{-K/2}, \tag{3.12}$$

with probability going to 1 as $p \to \infty$.

Remark 3.13 (Comparison to previous results). For fixed SNR, our initialization error rate with $K = 2$ agrees with the initialization error rate $\mathcal{O}(p^{-1})$ in matrix models (Gao et al., 2018). Furthermore, in the special case of non-degree TBMs with $\boldsymbol{\theta} = \mathbf{1}_p$, we achieve the same initial misclustering error $\mathcal{O}(p^{-K/2})$ as in non-degree models (Han et al., 2022a). Theorem 3.12 implies the advantage of our algorithm in achieving both accuracy and model flexibility.

Remark 3.14 (Failure of conventional tensor HOSVD). If we use conventional HOSVD for tensor denoising; that is, we use $\boldsymbol{U}_{\text{pre},k}$ in place of $\hat{\boldsymbol{U}}_k$ in line 2, then the misclustering rate becomes $\mathcal{O}(p^{-1})$ for all $K \geq 2$. This rate is substantially worse than our current rate (3.12).

Remark 3.15 (Singular-value gap-free clustering). Note that our clustering directly applies to the estimated mean tensor $\hat{\mathcal{X}}$ rather than the leading tensor factors $\hat{\boldsymbol{U}}_k$.

Applying clustering to the tensor factors suffers from the non-identifiability issue due to the infinitely many orthogonal rotations when the number of blocks $r \geq 3$ in the absence of singular-value gaps. Such ambiguity causes the trouble for effective clustering (Abbe et al., 2020). In contrast, our initialization algorithm applies the clustering to the overall mean tensor $\hat{\mathcal{X}}$. This strategy avoids the non-identifiability issue regardless of the number of blocks and singular-value gaps.

Algorithm 3 Multiway spherical clustering for dTBM

Sub-algorithm 1: Weighted higher-order initialization

Input: Observation $\mathcal{Y} \in \mathbb{R}^{p\times\cdots\times p}$, cluster number r, relaxation factor $\eta > 1$ in k-means clustering.

1: Compute factor matrices $\boldsymbol{U}_{\text{pre},k} = \text{SVD}_r(\text{Mat}_k(\mathcal{Y})), k \in [K]$ and the $(K-1)$-mode projections

$$\mathcal{X}_{\text{pre},k} = \mathcal{Y} \times_1 \boldsymbol{U}_{\text{pre},1}\boldsymbol{U}_{\text{pre},1}^T \times_2 \cdots \times_{k-1} \boldsymbol{U}_{\text{pre},k-1}\boldsymbol{U}_{\text{pre},k-1}^T \times_{k+1} \boldsymbol{U}_{\text{pre},k+1}\boldsymbol{U}_{\text{pre},k+1}^T \times_{k+2} \cdots \times_K \boldsymbol{U}_{\text{pre},K}\boldsymbol{U}_{\text{pre},K}^T.$$

2: Compute factor matrices $\hat{\boldsymbol{U}}_k = \text{SVD}_r(\text{Mat}_k(\mathcal{X}_{\text{pre},k})), k \in [K]$ and the denoised tensor

$$\hat{\mathcal{X}} = \mathcal{Y} \times_1 \hat{\boldsymbol{U}}_1\hat{\boldsymbol{U}}_1^T \times_2 \cdots \times_K \hat{\boldsymbol{U}}_K\hat{\boldsymbol{U}}_K^T.$$

3: **for** $k \in [K]$ **do**

4: Let $\hat{\boldsymbol{X}} = \text{Mat}_k(\hat{\mathcal{X}})$ and $S_0 = \{i \in [p] : \|\hat{\boldsymbol{X}}_{i:}\| = 0\}$. Set $\hat{z}(i)$ randomly in $[r]$ for $i \in S_0$.

5: For all $i \in S_0^c$, compute normalized rows $\hat{\boldsymbol{X}}^s_{i:} := \|\hat{\boldsymbol{X}}_{i:}\|^{-1}\hat{\boldsymbol{X}}_{i:}$.

6: Solve the clustering $\hat{z}_k \colon [p] \to [r]$ and centroids $\{\hat{\boldsymbol{x}}_j\}_{j\in[r]}$ using weighted k-means, such that

$$\sum_{i\in S_0^c} \|\hat{\boldsymbol{X}}_{i:}\|^2\|\hat{\boldsymbol{X}}^s_{i:} - \hat{\boldsymbol{x}}_{\hat{z}_k(i)}\|^2 \le \eta \min_{\bar{\boldsymbol{x}}_j, j\in[r], \bar{z}_k(i), i\in S_0^c} \sum_{i\in S^c} \|\hat{\boldsymbol{X}}_{i:}\|^2\|\hat{\boldsymbol{X}}^s_{i:} - \bar{\boldsymbol{x}}_{\bar{z}_k(i)}\|^2.$$

7: **end for**

Output: Initial clustering $z_k^{(0)} \leftarrow \hat{z}_k, k \in [K]$.

Sub-algorithm 2: Angle-based iteration

Input: Observation $\mathcal{Y} \in \mathbb{R}^{p\times\cdots\times p}$, initialization $z_k^{(0)} \colon [p] \to [r], k \in [K]$ from Sub-algorithm 1, iteration number T.

8: **for** $t = 0$ to $T-1$ **do**

9: Update the block tensor $\mathcal{S}^{(t)}$ via $\mathcal{S}^{(t)}(a_1, ..., a_K) = \text{Ave}\{\mathcal{Y}(i_1, \ldots, i_K) : z_k^{(t)}(i_k) = a_k, k \in [K]\}$.

10: **for** $k \in [K]$ **do**

11: Calculate the reduced tensor $\mathcal{Y}_k^{\text{d}} \in \mathbb{R}^{r\times\cdots\times r\times p\times r\times\cdots\times r}$ via

$$\mathcal{Y}_k^{\text{d}}(a_1, \ldots, a_{k-1}, i, a_{k+1}, \ldots, a_K) = \text{Ave}\{\mathcal{Y}(i_1, \ldots, i_{k-1}, i, i_{k+1}, \ldots, i_K) : z^{(t)}(i_j) = a_j, j \ne k\}$$

12: Let $\boldsymbol{Y}_k^{\text{d}} = \text{Mat}_k(\mathcal{Y}^{\text{d}})$ and $J_0 = \{i \in [p] : \left\|\boldsymbol{Y}^{\text{d}}_{i:}\right\| = 0\}$. Set $z_k^{(t+1)}(i)$ randomly in $[r]$ for $i \in J_0$.

13: Let $\boldsymbol{S}_k^{(t)} = \text{Mat}_k(\mathcal{S}^{(t)})$. For all $i \in J_0^c$, update the cluster assignment by

$$z(i)_k^{(t+1)} = \arg\max_{a\in[r]} \cos\left(\boldsymbol{Y}^{\text{d}}_{k,i:},\ \boldsymbol{S}^{(t)}_{k,a:}\right).$$

14: **end for**

15: **end for**

Output: Estimated clustering $z_k^{(T)} : [p] \mapsto [r], k \in [K]$.

Angle-based iteration

Our Theorem 3.12 has shown the polynomially decaying error rate from our initialization. Now we improve the error rate to exponential decay using local iterations. We propose an angle-based local iteration to improve the outputs from Sub-algorithm 1. To gain the intuition, consider an one-dimensional degree-corrected clustering problem with data vectors $\boldsymbol{x}_i = \theta(i)\boldsymbol{s}_{z(i)} + \boldsymbol{\epsilon}_i, i \in [p]$, where $\boldsymbol{s}_i$'s are known cluster centroids, $\theta(i)$'s are unknown positive degrees, and $z\colon [p] \mapsto [r]$ is the cluster assignment of interest. The angle-based k-means algorithm estimates the assignment z by minimizing the angle between data vectors and centroids; i.e.,

$$z(i) = \arg\max_{a\in[r]} \cos(\boldsymbol{x}_i,\ \boldsymbol{s}_a), \quad \text{for all } i \in [p]. \tag{3.13}$$

The classical Euclidean-distance based clustering (Han et al., 2022a) fails to recover z in the presence of degree heterogeneity, even under noiseless case. In contrast, the proposed angle-based k-means algorithm achieves accurate recovery without the explicit estimation of $\boldsymbol{\theta}$.

Our Sub-algorithm 2 shares the same spirit as in the angle-based k-means. We still take the order-3 tensor for illustration. Specifically, Sub-algorithm 2 updates estimated core tensor and cluster assignment in each iteration. We use superscript $\cdot^{(t)}$ to denote the estimate from the t-th iteration, where $t = 1, 2, \ldots$. For core tensor, we consider the following update strategy

$$\mathcal{S}^{(t)}(a_1, a_2, a_3) = \text{Ave}\{\mathcal{Y}(i_1, i_2, i_3)\colon z_k^{(t)}(i_k) = a_k, k \in [3]\}.$$

Intuitively, $\mathcal{S}^{(t)}$ becomes closer to the true core $\mathcal{S}$ as $z_k^{(t)}$ is more precise. For cluster assignment, we first aggregate the slices of $\mathcal{Y}$ and obtain the reduced tensor $\mathcal{Y}_1^{\text{d}} \in \mathbb{R}^{p\times r\times r}$ on the first mode with given $z_k^{(t)}$, where

$$\mathcal{Y}_1^{\text{d}}(i, a_2, a_3) = \text{Ave}\{\mathcal{Y}(i, i_2, i_3)\colon z_k^{(t)}(i_k) = a_k, k \neq 1\}.$$

Similarly, we also obtain $\mathcal{Y}_2^{\text{d}}, \mathcal{Y}_3^{\text{d}}$. We use $\boldsymbol{Y}_k^d$ and $\boldsymbol{S}_k^{(t)}$ to denote the $\text{Mat}_k(\mathcal{Y}^{\text{d}})$

and $\text{Mat}_k(\mathcal{S}^{(t)})$. The rows $\boldsymbol{Y}^d_{k,i:}$ and $\boldsymbol{S}^{(t)}_{k,a:}$ correspond to the $\boldsymbol{x}_i$ and $\boldsymbol{s}_a$ in the one-dimensional clustering (3.13). Then, we obtain the updated assignment by

$$z_k(i)^{(t+1)} = \arg\max_{a\in[r]} \cos\left(\boldsymbol{Y}^{\mathrm{d}}_{k,i:}, \boldsymbol{S}^{(t)}_{k,a:}\right), \quad \text{for all } i \in [p],$$

provided that $\boldsymbol{S}^{(t)}_{k,a:}$ is a non-zero vector. Otherwise, if $\boldsymbol{S}^{(t)}_{k,a:}$ is a zero vector, then we make the convention to assign $z_k^{(t+1)}(i)$ randomly in $[r]$. Full procedures for our angle-based iteration are described in Sub-algorithm 2.

We now establish the misclustering error rate of iterations under the stability assumption.

Definition 3.16 (Locally linear stability)**.** *Define the ε-neighborhood of z by $\mathcal{N}(z,\epsilon) = \{\bar{z}\colon \ell(\bar{z},z) \le \epsilon\}$. Let $\bar{z}\colon [p]\to[r]$ be a clustering function. We define two vectors associated with $\bar{z}$,*

$$\boldsymbol{p}(\bar{z}) = (|\bar{z}^{-1}(1)|, \ldots, |\bar{z}^{-1}(r)|)^T,$$
$$\boldsymbol{p}_{\boldsymbol{\theta}}(\bar{z}) = (\|\boldsymbol{\theta}_{\bar{z}^{-1}(1)}\|_1, \ldots, \|\boldsymbol{\theta}_{\bar{z}^{-1}(r)}\|_1)^T.$$

We call the degree is ε-locally linearly stable if and only if

$$\sin(\boldsymbol{p}(\bar{z}),\ \boldsymbol{p}_{\boldsymbol{\theta}}(\bar{z})) \lesssim \varepsilon\Delta_{\min}, \quad \textit{for all } \bar{z}\in\mathcal{N}(z,\varepsilon). \tag{3.14}$$

Roughly speaking, the vector $\boldsymbol{p}(\bar{z})$ represents the raw cluster sizes, and $\boldsymbol{p}_\theta(\bar{z})$ represents the relative cluster sizes weighted by degrees. The local stability holds trivially for $\varepsilon = 0$ based on the construction of parameter space (3.2). The condition (3.14) controls the impact of node degree to the $\boldsymbol{p}_\theta(\cdot)$ with respect to the misclustering rate ε and angle gap. Intuitively, the condition (3.14) controls the skewness of degree so that the angle between raw cluster size and degree-weighted cluster size is well controlled. The stability assumption is proposed for technical convenience, and we relax this condition in numerical studies; see Section 3.6.

Theorem 3.17 (Error for angle-based iteration)**.** *Consider the general sub-Gaussian dTBM with fixed $r \geq 1$, $K \geq 2$, independent noise under the parameter space $\mathcal{P}$, and Assumption 3.2. Assume that the locally linear stability of degree holds in the neighborhood $\mathcal{N}(z, \varepsilon)$ for all $\varepsilon \leq E_0$ and some $E_0 \gtrsim \log^{-1} p$. Let $\{z_k^{(0)}\}_{k=1}^K$ be the initialization for Sub-algorithm 2 and $z_k^{(t)}$ be the t-th iteration output on the k-th mode. Suppose $\min_{i\in[p]} \theta(i) \geq c$ for some constant $c > 0$, the SNR $\geq \tilde{C} p^{-(K-1)} \log p$ for some sufficiently large positive constant $\tilde{C}$, and the initialization satisfies*

$$L(z_k^{(0)}, z) \lesssim \frac{\Delta_{\min}^2}{r \log p}, \quad k \in [K].$$

With probability going to 1 as $p \to \infty$, there exists a contraction parameter $\rho \in (0, 1)$ such that

$$\ell(z, \hat{z}_k^{(t+1)}) \lesssim \underbrace{SNR^{-1} \exp\left(-\frac{p^{K-1} SNR}{r^{K-1}}\right)}_{\text{statistical error}} + \underbrace{\rho^t \ell(z, z_k^{(0)})}_{\text{computational error}}. \tag{3.15}$$

From the conclusion (3.15), we find that the iteration error is decomposed into two parts: statistical error and computational error. The statistical error is unavoidable with noisy data regardless t, whereas the computational error decays in an exponential rate as the number of iterations $t \to \infty$.

Corollary 3.18 (Exact recovery of dTBM with weighted higher-order initialization)**.** *Let the initialization $\{z_k^{(0)}\}_{k=1}^K$ be the output from Sub-algorithm 1. Assume SNR $\gtrsim p^{-K/2} \log p$. Combining all parameter assumptions and the results in Theorems 3.12 and 3.17, with probability going to 1 as $p \to \infty$, our estimate $z_k^{(T)}$ achieves exact recovery within polynomial iterations; more precisely,*

$$z_k^{(T)} = \pi_k \circ z, \quad \textit{for all } T \gtrsim \log_{1/\rho} p \textit{ and } k \in [K].$$

for some permutation $\pi_k \in \Pi$.

Therefore, our combined algorithm is *computationally efficient* as long as SNR $\gtrsim p^{-K/2} \log p$. Note that, ignoring the logarithmic term, the minimal SNR requirement,

$p^{-K/2}$, coincides with the computational critical value in Theorem 3.9. Therefore, our algorithm is optimal regarding the signal requirement and lies in the sharpest *computationally efficient* regime in Figure 3.2.

Extension to Bernoulli observations

Bernoulli or network observations are common in multiple fields. Our iteration Theorem 3.17 holds for Bernoulli models, but our initialization Theorem 3.12 does not. Moreover, our current dTBM is insufficient to address sparsity with decaying mean tensor. Here, we provide extra discussions for Bernoulli initialization and strategies under sparse settings.

- *Extension to dense binary dTBMs.* The main difficulty to establish initialization guarantees for Bernoulli observations lies in the denoising step (lines 1-2 in Sub-algorithm 1). We now provide a high-level explanation for the technical difficulty when applying Theorem 3.12 to Bernoulli observations.

  The derivation of Theorem 3.12 relies on the upper bound of the estimation error for the mean tensor in Lemma B.7; i.e., with high probability

  $$\|\hat{\mathcal{X}} - \mathcal{X}\|_F^2 \lesssim p^{K/2}, \tag{3.16}$$

  where $\mathcal{X} = \mathbb{E}\mathcal{Y}$ and $\hat{\mathcal{X}}$ is defined in Step 2 of Sub-algorithm 1. Unfortunately, the inequality (3.16) holds only for i.i.d. sub-Gaussian observations, while Bernoulli observations are generally not identically distributed.

  One possible remedy is to apply singular value decomposition to the *square unfolding* (Mu et al., 2014), $\mathrm{Mat}_{sq}(\cdot)$, of Bernoulli tensor $\mathcal{Y} \in \{0,1\}^{p_1 \times \cdots \times p_K}$. Specifically, the square matricization $\mathrm{Mat}_{sq}(\mathcal{Y}) \in \{0,1\}^{p^{\lfloor K/2 \rfloor} \times p^{\lceil K/2 \rceil}}$ has entries

$[\text{Mat}_{sq}(\mathcal{Y})](j_1, j_2) = \mathcal{Y}(i_1, \ldots, i_K)$, where

$$\begin{aligned} j_1 &= i_1 + p_1(i_2 - 1) + \cdots + p_1 \cdots p_{\lfloor K/2 \rfloor - 1}(i_{\lfloor K/2 \rfloor} - 1), \\ j_2 &= i_{\lceil K/2 \rceil} + p_{\lceil K/2 \rceil}(i_{\lceil K/2 \rceil + 1} - 1) + \cdots \\ &\qquad + p_{\lceil K/2 \rceil} \cdot p_{K-1}(i_K - 1). \end{aligned}$$

The matrix $\text{Mat}_{sq}(\mathcal{Y})$ is asymmetric. We interpret $\text{Mat}_{sq}(\mathcal{Y})$ as the adjacency matrix for a bipartite network with connections between two groups of nodes. The two groups of nodes in the bipartite network have $p_1 \cdots p_{\lfloor K/2 \rfloor}$ and $p_{\lceil K/2 \rceil} \cdots p_K$ nodes, respectively. The entry $[\text{Mat}_{sq}(\mathcal{Y})](j_1, j_2)$ refers to the presence of connection between the nodes indexed by combinations $(i_1, \ldots, i_{\lfloor K/2 \rfloor})$ and $(i_{\lceil K/2 \rceil}, \ldots, i_K)$. We summarize the procedure in Sub-algorithm 3.

Sub-algorithm 3: Weighted higher-order initialization for Bernoulli observation

Input: Bernoulli tensor $\mathcal{Y} \in \{0, 1\}^{p \times \cdots \times p}$, cluster number r, relaxation factor $\eta > 1$ in k-means clustering.

1: Let the matrix $\text{Mat}_{sq}(\mathcal{Y}) \in \{0, 1\}^{p^{\lfloor K/2 \rfloor} \times p^{\lceil K/2 \rceil}}$ denote the nearly square unfolded tensor. Compute the estimate $\mathcal{X}'$, where

$$\hat{\mathcal{X}}' = \underset{\text{rank}(\text{Mat}_{sq}(\mathcal{X})) \le r^{\lceil K/2 \rceil}}{\arg\min} \|\text{Mat}_{sq}(\mathcal{X}) - \text{Mat}_{sq}(\mathcal{Y})\|_F^2. \tag{3.17}$$

2: Implement lines 3-5 of Sub-algorithm 1 with $\hat{\mathcal{X}}$ replaced by $\hat{\mathcal{X}}'$ in (3.17).

Output: Initial clustering $z_k^{(0)} \leftarrow \hat{z}_k, k \in [K]$.

Proposition 3.5 (Error for Bernoulli initialization)**.** *Consider the Bernoulli dTBM in the parameter space $\mathcal{P}$ with fixed $r \ge 1, K \ge 2$. Assume that Assumption 3.2 holds, $\boldsymbol{\theta}$ is balanced, and $\min_{i \in [p]} \theta(i) \ge c$ for some constant $c > 0$. Let $z_k^{(0)}$ denote the output of Sub-algorithm 3. With probability going to 1 as $p \to \infty$, we have*

$$\ell(z_k^{(0)}, z_k) \lesssim \frac{r^K p^{-\lfloor K/2 \rfloor}}{SNR}, \quad \textit{and} \quad L(z_k^{(0)}, z_k) \lesssim \sigma^2 r^K p^{-\lfloor K/2 \rfloor}.$$

Remark 3.19 (Comparison with Gaussian model). The Bernoulli bound $\mathcal{O}(p^{-\lfloor K/2 \rfloor})$

in Proposition 3.5 is relatively looser than the Gaussian bound $\mathcal{O}(p^{-K/2})$ in Theorem 3.12. The gap between Bernoulli and Gaussian error decreases as the order K increases. Nevertheless, combining with angle iteration Sub-algorithm 2, Bernoulli clustering still achieves exponential error rate $\exp\left(-p^{(K-1)}\right)$ at a price of a larger SNR. The investigation of the gap between upper bound $p^{-\lfloor K/2 \rfloor}$ and the lower bound $p^{-K/2}$ for Bernoulli tensors will be left as future work. In numerical experiments, we will use our original initialization, Sub-algorithm 1, to verify the robustness to Bernoulli observations.

Remark 3.20 (Comparison with previous methods). Previous work (Ke et al., 2019) develops a spectral clustering method for Bernoulli dTBM. Ke et al. (2019) adopts a different signal notion based on the singular gap in the core tensor, denoted as Δ_{singular}. By (Ke et al., 2019, Theorem 1), the spectral method achieves exact recovery with $\Delta_{\text{singular}} \gtrsim p^{-1/2}$. However, we are not able to infer the exact recovery of spectral method by our angle-base SNR condition. Consider an order-2 dTBM with $p > 2, \sigma^2 = 1$, $\boldsymbol{\theta} = \mathbf{1}_p$, equal size assignment $|z^{-1}(a)| = p/r$ for all $a \in [r]$, and core matrix equal to the 2-dimensional identity matrix $\boldsymbol{S} = \boldsymbol{I}_2$. The singular gap under this setting is $\Delta_{\text{singular}} = \min\{\lambda_1 - \lambda_2, \lambda_2\} = 0$, where $\lambda_1 \geq \lambda_2$ are singular values of $\boldsymbol{S}$. In contrast, our angle gap $\Delta_{\min}^2 = 2$ satisfies the SNR condition in Theorem 3.17. Then, our algorithm achieves the exact recovery, but the spectral method in Ke et al. (2019) fails.

Hence, for fair comparison, we compare the best performance of our algorithm and Ke et al. (2019) under the strongest signal setting of each model. Since both methods contain an iteration procedure, we set the iteration number to infinity to avoid the computational error. Considering the largest angle-based SNR $\asymp 1$ in Theorem 3.17, our Bernoulli clustering achieves exponential error rate of order $\exp(-p^{(K-1)})$; considering the largest singular gap $\Delta_{\text{singular}} \asymp 1$ in Theorem 1 of Ke et al. (2019), the spectral clustering has a polynomial error rate of order p^{-2}. Our algorithm still shows a better theoretical accuracy than the competitive work for Bernoulli observations.

- *Extension to sparse binary dTBMs.* The sparsity is often a popular feature in hypergraphs (Florescu and Perkins, 2016; Ke et al., 2019; Ahn et al., 2018). Specifically, the sparse binary dTBM assumes that, the entries of $\mathcal{Y}$ follow independent Bernoulli distributions with the mean

$$\mathbb{E}\mathcal{Y} = \alpha_p \mathcal{S} \times_1 \boldsymbol{\Theta}\boldsymbol{M} \times_2 \cdots \times_K \boldsymbol{\Theta}\boldsymbol{M}, \tag{3.18}$$

  where the extra scalar parameter $\alpha_p \in (0, 1]$ is function of p that controls the sparsity. A smaller α_p indicates a higher level of sparsity. Our current work focuses on dense dTBM with $\alpha_p = 1$. While sparse dTBM is an interesting application, the algorithm and its analysis require different techniques. Below, we discuss possible modifications of the algorithm.

  The sparsity affects our initialization guarantee in our Theorem 3.12. In our initialization, the spectral denoising step (lines 1-2 in Sub-algorithm 1) implements matrix SVD to unfolded tensors. However, SVD-based methods are believed to fail in extremely sparse SBM due to the localization phenomenon in the singular vectors (Florescu and Perkins, 2016). Inspired by Florescu and Perkins (2016), we adopt the diagonal-deleted HOSVD (D-HOSVD) (Ke et al., 2019) as the initialization in our higher-order clustering.

  The sparsity also affects the iteration guarantee in our Theorem 3.17. The decaying mean tensor leads to a worse statistical error of order $\mathcal{O}(-\alpha_p p^{K-1})$ on $\hat{\mathcal{X}}$. The theoretical analyses for sparse binary dTBM and algorithms are left as future directions. Instead, we add numerical experiments to evaluate the robustness of our algorithm and the improvement of D-HOSVD initialization in the sparse dTBM; see Appendix A.

Practical issues

Computational complexity. Our two-stage algorithm has a computational cost polynomial in tensor dimension p. Specifically, the complexity of Sub-algorithm 1 is $\mathcal{O}(Kp^{K+1} + Krp^K)$, where the first term is contributed by the double projection

and the calculation of $\hat{\mathcal{X}}$, and the second term comes from normalization and the k-means. The cost of each update in Sub-algorithm 2 is $\mathcal{O}(p^K + pr^K)$, where p^K comes from the calculation of $\mathcal{S}^{(t)}$ and $\mathcal{Y}_k^{\mathrm{d}}$, and pr^K comes from the normalization of $\mathcal{Y}_k^{\mathrm{d}}$, the calculation of $\mathcal{S}^{(t)}$, and the cluster assignment update in Step 13.

Hyper-parameter selection. In our theoretical analysis, we have assumed the true cluster number r is given to our algorithm. In practice, the cluster number r is often unknown, and we now propose a method to choose r from data. We impose the Bayesian information criterion (BIC) and choose the cluster number that minimizes BIC; i.e., under the symmetric Gaussian dTBM (3.1),

$$\hat{r} = \underset{r \in \mathbb{Z}_+}{\arg\min} \left(p^K \log(\|\hat{\mathcal{X}} - \mathcal{Y}\|_F^2) + p_e(r) K \log p \right), \tag{3.19}$$

with $\hat{\mathcal{X}} = \hat{\mathcal{S}}(r) \times_1 \hat{\boldsymbol{\Theta}}(r)\hat{\boldsymbol{M}}(r) \times_2 \cdots \times_K \hat{\boldsymbol{\Theta}}(r)\hat{\boldsymbol{M}}(r)$, where the triplet $(\hat{z}(r), \hat{\mathcal{S}}(r), \hat{\boldsymbol{\theta}}(r))$ are estimated parameters with cluster number r, and $p_e(r) = r^K + p(\log r + 1) - r$ is the effective number of parameters. Note that we have added the argument (r) to related quantities as functions of r. In particular, the estimate $\hat{\boldsymbol{\theta}}(r)$ in (3.19) is obtained by first calculating the reduced tensor $\hat{\mathcal{Y}}^{\mathrm{d}}$ with $\hat{z}(r)$, and then normalizing the row norms $\|\hat{\boldsymbol{Y}}_{i:}^{\mathrm{d}}\|$ to 1 in each cluster; i.e.,

$$\hat{\boldsymbol{\theta}}(r) = (\hat{\theta}(1, r), \ldots, \hat{\theta}(p, r))^T,$$

with $\hat{\theta}(i, r) = \|\hat{\boldsymbol{Y}}^{\mathrm{d}}(r)_{i:}\| / \sum_{j:\hat{z}(j,r)=\hat{z}(i,r)} \|\hat{\boldsymbol{Y}}^{\mathrm{d}}(r)_{j:}\|$, $\hat{\boldsymbol{Y}}^{\mathrm{d}}(r) = \mathrm{Mat}(\hat{\mathcal{Y}}^{\mathrm{d}}(r))$, $\hat{\mathcal{Y}}^{\mathrm{d}}(r)(i, a_2, \ldots, a_K) = \mathrm{Ave}\{\mathcal{Y}(i, i_2, \ldots, i_K) : \hat{z}(i_k, r) = a_k, k \neq 1\}$, and $\hat{z}(i, r)$ denotes the community label for the i-th node with given cluster number r. We evaluate the performance of the BIC criterion in Section 3.6.

3.5 Comparison with Non-degree Tensor Block Model

We discuss the connections and differences between dTBM and TBM (Han et al., 2022a) from three aspects: signal notions, theoretical results, and algorithms. Without loss of generality, let $\sigma^2 = 1$.

- *Signal notion.* The signal levels in both TBM (Han et al., 2022a) and our dTBM are functions of the core tensor $\mathcal{S}$. We emphasize that the signal notions are different between the two models. In particular, the Euclidean-based signal notion in TBM Han et al. (2022a) fails to accurately describe the phase transition in our dTBM due to the possible heterogeneity in degree $\boldsymbol{\theta}$. To compare, we denote our angle-based signal notion in (3.4) and the Euclidean-based SNR in Han et al. (2022a) as Δ^2_{ang} and Δ^2_{Euc}, respectively:

$$\Delta^2_{\text{ang}} = 2(1 - \max_{a \neq b \in [r]} \cos(\boldsymbol{S}_{a:}, \boldsymbol{S}_{b:})),$$
$$\Delta^2_{\text{Euc}} = \min_{a \neq b \in [r]} \|\boldsymbol{S}_{a:} - \boldsymbol{S}_{b:}\|^2.$$

  By Lemma B.3 in the Appendix B, we have

$$\Delta^2_{\text{ang}} \max_{a \in [r]} \|\boldsymbol{S}_{a:}\|^2 \leq \Delta^2_{\text{Euc}}.$$

  The above inequality indicates that the condition $\Delta^2_{\text{Euc}} \leq p^\gamma$ is sufficient but not necessary for $\Delta^2_{\text{ang}} \leq p^\gamma$. In fact, if we were to use Δ^2_{Euc} for both models, then the phase transition of dTBM can be arbitrarily worse than that for TBM.

  Here, we provide an example to illustrate the dramatical difference between TBM and dTBM with the same core tensor.

  Example 3.6 (Comparison with Euclidean-based signal notion)**.** *Consider a biclustering model with $\boldsymbol{\theta} = \mathbf{1}$ and an order-2 core matrix*

$$\boldsymbol{S} = \begin{pmatrix} p^{(\gamma+1)/2} + 2 & 2p^{(\gamma+1)/2} + 4 \\ 2 & 4 \end{pmatrix}, \quad \textit{with } \gamma \leq -1.$$

  The core matrix $\boldsymbol{S}$ lies in the parameter spaces of TBM and our dTBM. Here, the constraint $\gamma \leq -1$ is added to ensure the bounded condition of $\boldsymbol{S}$ in our parameter space in (3.2)*. The angle-based and Euclidean-based signal levels of $\boldsymbol{S}$ are*

$$\Delta^2_{ang}(\boldsymbol{S}) = 0 \ (\leq p^\gamma), \quad \Delta^2_{Euc}(\boldsymbol{S}) = 5p^{\gamma+1} \ (\geq p^\gamma).$$

We conclude that TBM with $\boldsymbol{S}$ achieves exact recovery with a polynomial-time algorithm; see (Han et al., 2022a, Theorem 4). By contrast, the dTBM with the same $\boldsymbol{S}$ and input $r = 2$ violets the identifiability condition, and thus fails to be solved by all estimators; see our Theorem 3.3.

- *Theoretical results.* In both works, we study the phase transition of TBM and dTBM with respect to the Euclidean and angle-based SNRs. We briefly summarize the results in Han et al. (2022a) and compare with ours.

  Statistical critical value:

$$\begin{aligned}
\text{Ours: } \Delta_{\text{ang}}^2 &\lesssim p^{-(K-1)} \Rightarrow \text{statistically impossible;} \\
\Delta_{\text{ang}}^2 &\gtrsim p^{-(K-1)} \Rightarrow \text{MLE achieves exact recovery;} \\
\text{Han's: } \Delta_{\text{Euc}}^2 &\lesssim p^{-(K-1)} \Rightarrow \text{statistically impossible;} \\
\Delta_{\text{Euc}}^2 &\gtrsim p^{-(K-1)} \Rightarrow \text{MLE achieves exact recovery.}
\end{aligned}$$

  Computational critical value:

$$\begin{aligned}
\text{Ours: } \Delta_{\text{ang}}^2 &\lesssim p^{-K/2} \Rightarrow \text{computationally impossible;} \\
\Delta_{\text{ang}}^2 &\gtrsim p^{-K/2} \Rightarrow \text{ computationally efficient;} \\
\text{Han's: } \Delta_{\text{Euc}}^2 &\lesssim p^{-K/2} \Rightarrow \text{computationally impossible;} \\
\Delta_{\text{Euc}}^2 &\gtrsim p^{-K/2} \Rightarrow \text{computationally efficient.}
\end{aligned}$$

  The above comparison reveals four major differences.

  First, none of our results in Section 3.3 are corollaries of Han et al. (2022a). Both models show the similar conclusion but under different conditions. While the TBM impossibility (Han et al., 2022a) provides a necessary condition for our dTBM impossibility, we find that such a condition is often loose. There exists a regime of $\mathcal{S}$ in which TBM problems are computationally efficient but dTBM problems are statistically impossible; see Example 3.6. This

observation has motivated us to develop the new signal notion Δ^2_{ang} for sharp dTBM phase transition conditions.

Second, to find the phase transition, we need to show both the impossibility and achievability when SNR is below and above the critical value, respectively. While the TBM impossibility can serve as a loose condition of our dTBM impossibility, more efforts are required to show the achievability. In particular, since TBM is a more restrictive model than dTBM, the achievability in Han et al. (2022a) does not imply the achievability of dTBM in a larger parameter space. The latter requires us to develop new MLE and polynomial algorithms for dTBM achievability.

Third, from the perspective of proofs, we develop new dTBM-specific techniques to handle the extra degree heterogeneity. In our Theorem 3.7, we construct a special non-trivial degree heterogeneity to establish the lower bound for arbitrary core tensor with small angle gap, while, TBM (Han et al., 2022a) considers the constructions without degree parameter. In our Theorem 3.9, we construct a rank-2 tensor to relate HPC conjecture to Δ^2_{ang}, while TBM (Han et al., 2022a) constructs a rank-1 tensor to relate HPC conjecture to Δ^2_{Euc}. The asymptotic non-equivalence between Δ^2_{ang} and Δ^2_{Euc} renders our proof technically more involved.

Last, we discuss the statistical impossibility statements. Our Theorem 3.7 implies the statistical impossibility whenever the core tensor $\mathcal{S}$ leads to an angle-based SNR below the critical value, while, Theorem 6 in Han et al. (2022a) implies the worst case statistical impossibility for a particular core tensor $\mathcal{S}$ with Euclidean-based SNR below the statistical limit. Hence, our Theorem 3.7 shows a stronger statistical impossibility for dTBM than that presented in TBM (Han et al., 2022a, Theorem 6). However, inspecting the proof of Han et al. (2022a), the proof of Theorem 6 indeed implies a stronger TBM impossibility statement for arbitrary core tensor; i.e., when $\gamma < -(K-1)$

$$\liminf_{p\to\infty} \inf_{\mathcal{S}\in\mathcal{P}_{\mathcal{S},\text{TBM}}\cap\{\Delta^2_{\text{Euc}}=p^\gamma\}} \inf_{\hat{z}_{\text{stats}}} \sup_{z\in\mathcal{P}_{z,\text{TBM}}} \mathbb{E}[p\ell(\hat{z}_{\text{stats}}, z)] \geq 1,$$

where $\mathcal{P}_{\mathcal{S},\text{TBM}}$ and $\mathcal{P}_{z,\text{TBM}}$ refer to the space for core tensor $\mathcal{S}$ and assignment z under TBM, respectively. Again, in terms of the strong statistical impossibility, both models show the similar conclusion but under different conditions. Since two impossibilities consider different core tensor regimes with non-equivalent Δ^2_{ang} and Δ^2_{Euc}, we emphasize that different proof techniques are required to obtain these similar conclusions. See our proof sketch in Section 3.8, Appendices B.2 and B.2 for detail technical differences.

- *Algorithms.* Both Han et al. (2022a) and our work propose the two-step algorithm, which combines warm initialization and iterative refinement to achieve exact recovery. This local-to-global strategy is not new in clustering literature (Gao and Zhang, 2022; Chien et al., 2019). The highlight of our algorithm is the angle-based update in lines 10-14, Sub-algorithm 2, which is specifically designed for dTBM to avoid the estimation of $\boldsymbol{\theta}$. This angle-based update brings new proof challenges. We develop polar-coordinate based techniques to establish the error rate for the proposed algorithm.

3.6 Numerical studies

We evaluate the performance of the weighted higher-order initialization and angle-based iteration in this section. We report average errors and standard deviations across 30 replications in each experiment. Clustering accuracy is assessed by clustering error rate (CER, i.e., one minus rand index). The CER between $(\hat{z}, z)$ is equivalent to misclustering error $\ell(\hat{z}, z)$ up to constant multiplications (Meilă, 2012), and a lower CER indicates a better performance.

We generate order-3 tensors with *assortative* (Gao et al., 2018) core tensors to control SNR; i.e., we set $\mathcal{S}_{aaa} = s_1$ for $a \in [r]$ and others be s_2, where $s_1 > s_2 > 0$. Let $\alpha = s_1/s_2$. We set α close to 1 such that $1 - \alpha = o(p)$. In particular, we have $\alpha = 1 + \Omega(p^{\gamma/2})$ with $\gamma < 0$ by Assumption 3.2 and definition (3.4). Hence, we easily adjust SNR via varying α. The assortative setting is proposed for simulations, and our algorithm is applicable for general tensors in practice. The cluster assignment

z is randomly generated with equal probability across r clusters for each mode. Without further explanation, we generate degree heterogeneity $\boldsymbol{\theta}$ from absolute normal distribution by $\theta(i) = |X_i| + 1 - 1/\sqrt{2\pi}$ with $|X_i| \overset{\text{i.i.d.}}{\sim} N(0,1), i \in [p]$ and normalize $\boldsymbol{\theta}$ to satisfy (3.2). Also, we set $\sigma^2 = 1$ for Gaussian data without further specification.

Verification of theoretical results

The first experiment verifies statistical-computational gap described in Section 3.3. Consider the Gaussian model with $p = \{80, 100\}$, $r = 5$. We vary γ in $[-1.2, -0.4]$ and $[-2.1, -1.4]$ for matrix $(K = 2)$ and tensor $(K = 3)$ clustering, respectively. Note that finding MLE under dTBM is computationally intractable. We approximate MLE using an oracle estimator, i.e., the output of Sub-algorithm 2 initialized from true assignment. Figure 3.4a shows that both our algorithm and oracle estimator start to decrease around the critical value $\gamma_{\text{stat}} = \gamma_{\text{comp}} = -1$ in matrix case. In contrast, Figure 3.4b shows a significant gap in the phase transitions between the algorithm estimator and oracle estimator in tensor case. The oracle error rapidly decreases to 0 when $\gamma_{\text{stat}} = -2$, whereas the algorithm estimator tends to achieve exact clustering when $\gamma_{\text{comp}} = -1.5$. Figure 3.4 confirms the existence of the statistical-computational gap in our Theorems 3.7 and 3.9.

The second experiment verifies the performance guarantees of two algorithms: (i) weighted higher-order initialization; (ii) combined algorithm of weighted higher-order initialization and angle-based iteration. We consider both the Gaussian and Bernoulli models with $p = \{80, 100\}$, $r = 5$, $\gamma \in [-2.1, -1.4]$. Figure 3.5 shows the substantial improvement of combined algorithm over initialization, especially under weak and intermediate signals. This phenomenon agrees with the error rates in Theorems 3.12 and 3.17 and confirms the necessity of the local iterations.

The third experiment evaluates the empirical performance of the BIC criterion to select unknown cluster number. We generate the data from an order-3 Gaussian model with $p = \{50, 80\}$, $r = \{2, 4\}$, and noise level $\sigma^2 \in \{0.25, 1\}$. Table 3.3 shows that our BIC criterion well chooses the true r under most settings. Note that the BIC

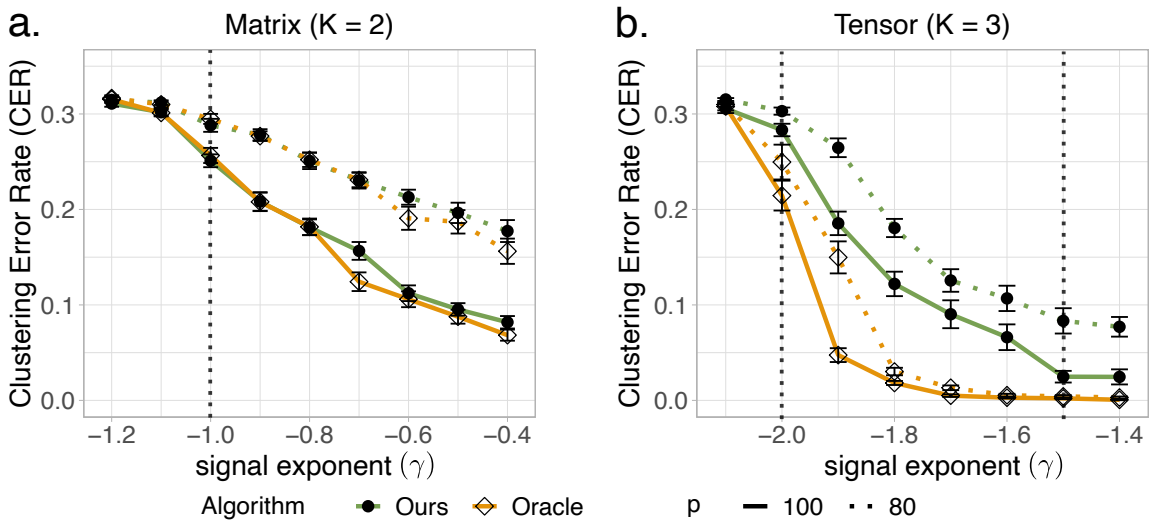

Figure 3.4: SNR phase transitions for clustering in dTBM with $p = \{80, 100\}, r = 5$ under (a) matrix case with $\gamma \in [-1.2, -0.4]$ and (b) tensor case with $\gamma \in [-2.1, -1.4]$.

Settings	$p = 50, \sigma^2 = 0.25$		$p = 50, \sigma^2 = 1$		$p = 80, \sigma^2 = 0.25$		$p = 80, \sigma^2 = 1$	
r	2	4	2	4	2	4	2	4
$\hat{r}$	2(0)	3.9(0.25)	2(0)	3.1(0.52)	2(0)	4(0)	2(0)	3.9(0.31)

Table 3.3: Estimated cluster number given by BIC criterion under the low noise level $(\sigma^2 = 0.25)$ and high noise level $(\sigma^2 = 0.5)$ settings. Numbers in parentheses are standard deviations of $\hat{r}$ over 30 replications.

slightly underestimates the true cluster number $(r = 4)$ with smaller dimension and higher noise $(p = 50, \sigma^2 = 1)$, and the accuracy immediately increases with larger dimension $p = 80$. The improvement follows from the fact that a larger dimension p indicates a larger sample size in the tensor block model. Therefore, we conclude that BIC criterion is a reasonable way to tune the cluster number.

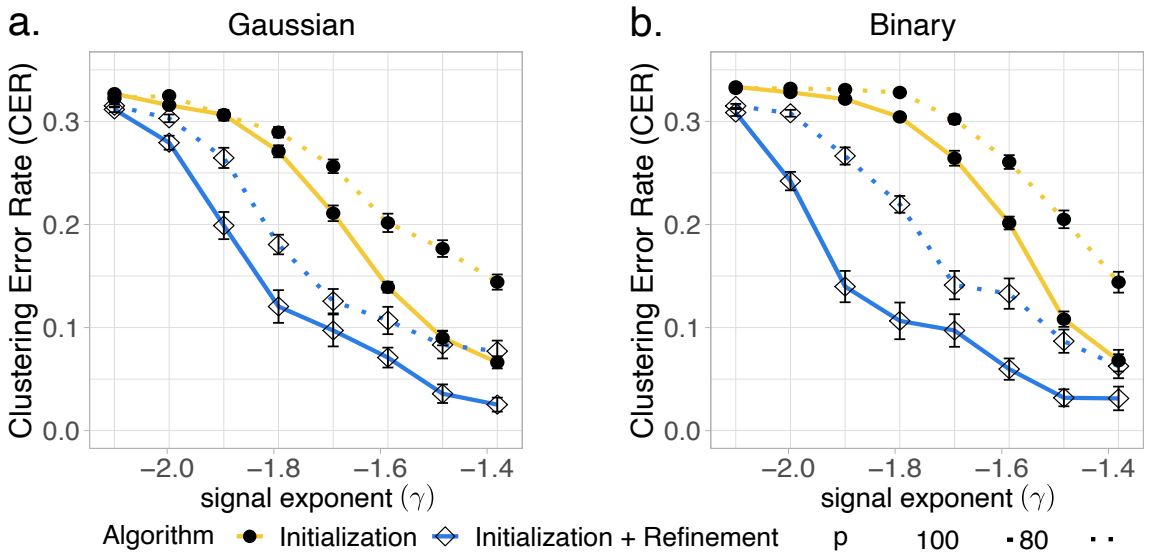

Figure 3.5: CER versus signal exponent (γ) for initialization only and for combined algorithm. We set $p = \{80, 100\}, r = 5, \gamma \in [-2.1, -1.4]$ under (a) Gaussian models and (b) Bernoulli models.

Comparison with other methods

We compare our algorithm with following higher-order clustering methods:

- **HOSVD**: HOSVD on data tensor and k-means on the rows of the factor matrix;
- **HOSVD+**: HOSVD on data tensor and k-means on the ℓ_2-normalized rows of the factor matrix;
- **HLloyd** (Han et al., 2022a): High-order clustering algorithm developed for non-degree tensor block models;
- **SCORE** (Ke et al., 2019): Tensor-SCORE for clustering developed for sparse binary tensors.

Among the four alternative algorithms, the **SCORE** is the closest method to ours. We set the tuning parameters of **SCORE** as in previous literature (Ke et al., 2019). The

methods **SCORE** and **HOSVD+** are designed for degree models, whereas **HOSVD** and **HLloyd** are designed for non-degree models. We conduct two experiments to assess the impacts of (i) signal strength and (ii) degree heterogeneity, based on Gaussian and Bernoulli models with $p = 100, r = 5$. We refer to our algorithm as **dTBM** in the comparison.

We investigate the effects of signal to clustering performance by varying $\gamma \in [-1.5, -1.1]$. Figure 3.6 shows that our method **dTBM** outperforms all other algorithms. The sub-optimality of **SCORE** and **HOSVD+** indicates the necessity of local iterations on the clustering. Furthermore, Figure 3.6 shows the inadequacy of non-degree algorithms in the presence of mild degree heterogeneity. The experiment demonstrates the benefits of addressing heterogeneity in higher-order clustering tasks.

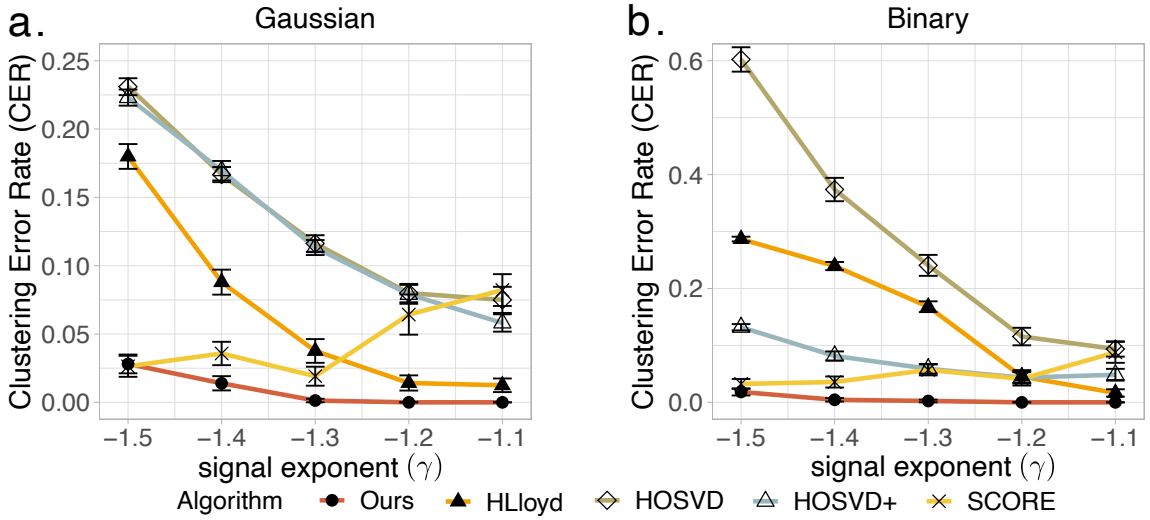

Figure 3.6: CER versus signal exponent (denoted γ) for different methods. We set $p = 100, r = 5, \gamma \in [-1.5, -1.1]$ under (a) Gaussian and (b) Bernoulli models.

The only exception in Figure 3.6 is the slightly better performance of **HLloyd** over **HOSVD+** under Gaussian model. However, we find the advantage of **HLloyd** disappears with higher degree heterogeneity. We perform extra simulations to

verify the impact of degree effects. We use the same setting as in the first experiment in the Section 3.6, except that we now generate the degree heterogeneity $\boldsymbol{\theta}$ from Pareto distribution prior to normalization. The density function of Pareto distribution is $f(x|a,b) = ab^a x^{-(a+1)}\mathbb{1}\{x \geq b\}$, where a is called *shape* parameter. We vary $a \in \{2, 6\}$ and choose b such that $\mathbb{E}X = a(a-1)^{-1}b = 1$ for X following Pareto(a, b). Note that a smaller a leads to a larger variance in $\boldsymbol{\theta}$ and hence a larger degree heterogeneity. We consider the Gaussian model under low $(a = 6)$ and high $(a = 2)$ degree heterogeneity. Figure 3.7 shows that the errors for non-degree algorithms (**HLloyd**, **HOSVD**) increase with degree heterogeneity. In addition, the advantage of **HLloyd** over **HOSVD+** disappears with higher degree heterogeneity.

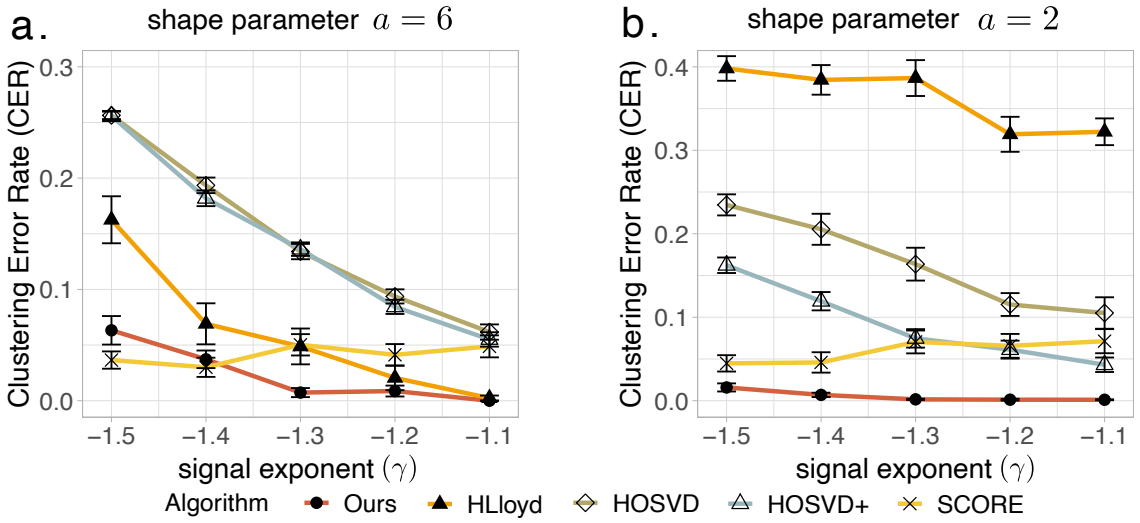

Figure 3.7: CER comparison versus signal exponent (denoted γ) under (a) low (shape parameter $a = 6$) (b) high (shape parameter $a = 2$) degree heterogeneity. We set $p = 100, r = 5, \gamma \in [-1.5, -1.1]$ under Gaussian model.

The last experiment investigates the effects of degree heterogeneity to clustering performance. We fix the signal exponent $\gamma = -1.2$ and vary the extent of degree heterogeneity. In this experiment, we generate $\boldsymbol{\theta}$ from Pareto distribution prior to normalization. We vary the shape parameter $a \in [3, 6]$ in the Pareto distribution to investigate a range of degree heterogeneities. Figure 3.8 demonstrates the stability

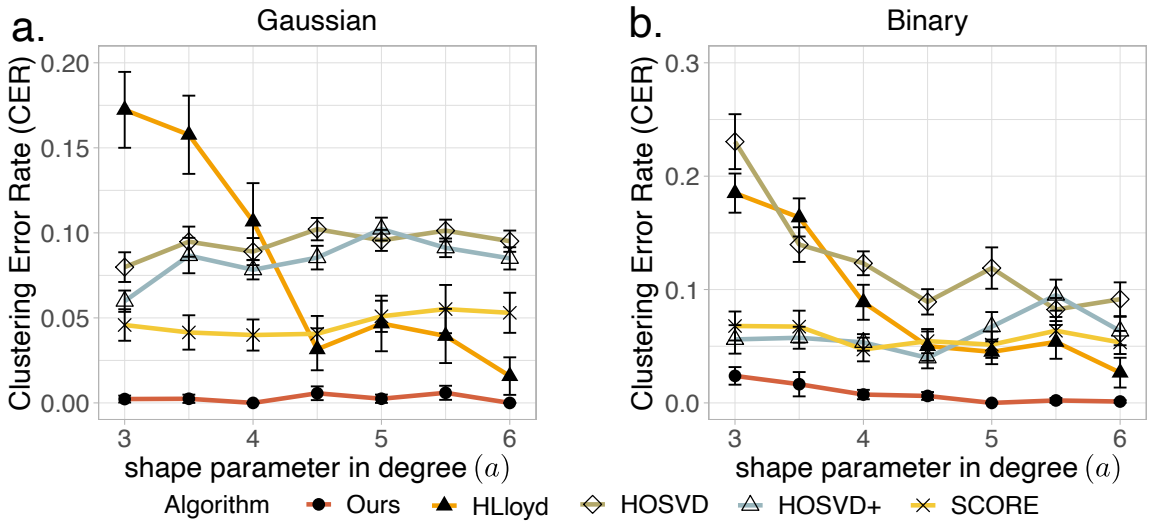

Figure 3.8: CER versus shape parameter in degree (denoted $a \in [3, 6]$) for different methods. We set $p = 100, r = 5, \gamma = -1.2$ under (a) Gaussian and (b) Bernoulli models.

of degree-corrected algorithms (**dTBM**, **SCORE**, **HOSVD+**) over the entire range of degree heterogeneity under consideration. In contrast, non-degree algorithms (**HLloyd**, **HOSVD**) show poor performance with large heterogeneity, especially in Bernoulli cases. This experiment, again, highlights the benefit of addressing degree heterogeneity in higher-order clustering.

3.7 Real data applications

Human brain connectome data analysis

The Human Connectome Project (HCP) aims to construct the structural and functional neural connections in human brains (Van Essen et al., 2013). We preprocess the original dataset following Desikan et al. (2006) and partition the brain into 68 regions. The cleaned dataset includes brain networks for 136 individuals. Each brain network is represented by a 68-by-68 binary symmetric matrix, where the

entry with value 1 indicates the presence of connection between node pairs, while the value 0 indicates the absence. We use $\mathcal{Y} \in \{0,1\}^{68\times68\times136}$ to denote the binary tensor. Individual attributes such as gender and sex are recorded.

We apply our general asymmetric algorithm to the HCP data with the numbers of clusters on three modes $r_1 = r_2 = 4$ and $r_3 = 3$. The selection of r_1 and r_2 follows the human brain anatomy and the symmetry in the brain network, and the r_3 is specified following previous analysis (Hu et al., 2022). Because of the symmetry in the data, the estimated brain node clustering results are the same on the first and second modes. Figure 3.9 shows that brain connection exhibits a strong spatial separation structure. Specifically, the first cluster, named *L.Hemis*, involves all the nodes in the left hemisphere. The nodes in the right hemisphere are further separated into three clusters led by the middle-part tissues in Temporal and Parietal lobes (*R.Temporal*), the back-part tissues in Occipital lobe (*R.Occipital*), and the front-part tissues in Frontal and Parietal lobes (*R.Supra*). This clustering result is reasonable since the left and right hemispheres often play different roles in human brains.

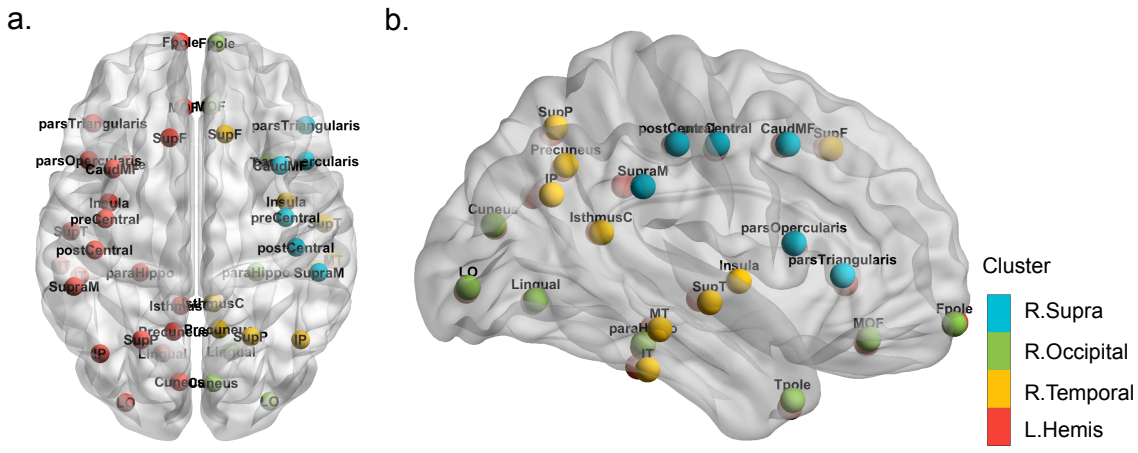

Figure 3.9: Illustration of brain node clustering results for HCP data with (a) top and (b) side views.

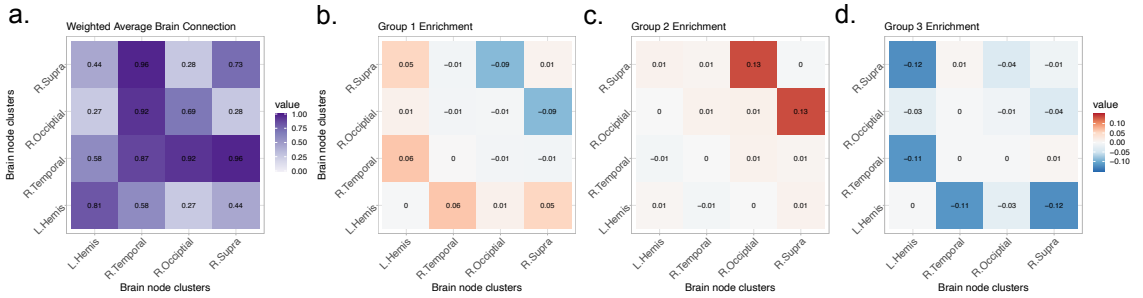

Figure 3.10: Mode 3 slices of estimated core tensor $\hat{\mathcal{S}}$. (a) Average estimated slice weighted by the group size; (b)-(d) Group-specified enrichment, i.e., the difference between each slice of $\hat{\mathcal{S}}$ and the averaged slice.

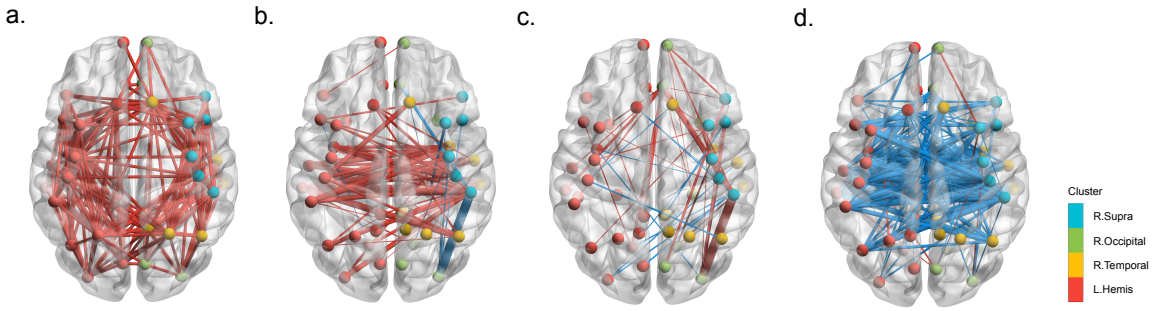

Figure 3.11: Observed brain connections in the population and each group of individuals. (a) Average brain network; (b)-(d) Group-specified brain network enrichments in Groups 1-3. Red edges represent the positive enrichment and blue edges represent the negative enrichment.

Figure 3.10 illustrates the estimated core tensor $\hat{\mathcal{S}}$ with estimated clustering, and Figure 3.11 visualizes the average brain connections and the connection enrichment in contrast to average networks in each group. In general, we find that the inner-hemisphere connection has stronger connection compared to inter-hemisphere connections (Figure 3.10a). Also, the back and front parts (*R.Occipital*, *R.Supra*) are shown to have more interactions with temporal tissues than inner-cluster connections. In addition, the group 1 with 54% females shows an enrichment on the

inter-hemisphere connections (Figure 3.10b), while group 4 with only 36% females exhibits a reduction (Figure 3.10d). This result agrees with previous findings in Hu et al. (2022). The enrichment on the back-front connection is also recognized in group 3 (Figure 3.10c). The interpretive patterns in our results demonstrate the usefulness of our clustering methods in the human brain connectome data application.

Peru Legislation data analysis

We also apply our method to the legislation networks in the Congress of the Republic of Peru (Lee et al., 2017). Because of the frequent political power shifts in the Peruvian Congress during 2006-2011, we choose to focus on the data for the first half of 2006-2007 year. The dataset records the co-sponsorship of 116 legislators from top 5 parties and 802 bill proposals. We reconstruct legislation network as an order-3 binary tensor $\mathcal{Y} \in \{0,1\}^{116\times116\times116}$, where $\mathcal{Y}_{ijk} = 1$ if the legislators (i, j, k) have sponsored the same bill, and $\mathcal{Y}_{ijk} = 0$ otherwise. The true party affiliations of legislators are provided and serve as the ground truth. We apply various higher-order clustering methods to $\mathcal{Y}$ with $r = 5$. Table 3.4 shows that our **dTBM** achieves the best performance compared to others. The second best method is the two-stage algorithm **HLloyd**, followed by the spectral methods **SCORE** and **HOSVD+**. This result is consistent with our simulations under strong signal and moderate degree heterogeneity. The comparison suggests that our method **dTBM** is more appealing in real-world applications.

Method	**dTBM**	**HOSVD**	**HOSVD+**	**HLloyd**	**SCORE**
CER	**0.116**	0.22	0.213	0.149	0.199

Table 3.4: Clustering errors (measured by CER) for various methods in the analysis of Peru Legislation dataset.

3.8 Proof Sketches

In this section, we provide the proof sketches for the main Theorem 3.7 (Impossibility), Theorem 3.9 (Impossibility), and Theorems 3.12-3.17. Detail proofs and extra theoretical results are provided in Appendix B.

Proof sketch of Theorem 3.7 (Impossibility) and Theorem 3.9 (Impossibility)

The proofs of impossibility in Theorems 3.7 and 3.9 share the same proof idea with (Han et al., 2022a, Theorems 6 and 7) and (Gao et al., 2018, Theorem 2). In both proofs of statistical and computational impossibilities, the key idea is to construct a particular set of parameters to lower bound the minimax rate. Specifically, for statistical impossibility in Theorem 3.7, we construct a particular $(z^*_{\text{stats}}, \boldsymbol{\theta}^*_{\text{stats}}) \in \mathcal{P}_{z,\boldsymbol{\theta}}$ such that for all $\mathcal{S}^* \in \mathcal{P}_{\mathcal{S}}(\gamma)$

$$\begin{aligned}&\inf_{\hat{z}_{\text{stats}}} \sup_{(z,\boldsymbol{\theta})\in\mathcal{P}_{z,\boldsymbol{\theta}}} \mathbb{E}[p\ell(\hat{z}_{\text{stat}}, z)] \\ &\geq \inf_{\hat{z}_{\text{stats}}} \mathbb{E}[p\ell(\hat{z}_{\text{stat}}, z^*_{\text{stats}})|(z^*_{\text{stats}}, \mathcal{S}^*, \boldsymbol{\theta}^*_{\text{stats}})] \geq 1;\end{aligned} \tag{3.20}$$

for computational impossibility in Theorem 3.9, we construct a particular $(z^*_{\text{comp}}, \mathcal{S}^*_{\text{comp}}, \boldsymbol{\theta}^*_{\text{comp}}) \in \mathcal{P}(\gamma)$ such that

$$\begin{aligned}&\inf_{\hat{z}_{\text{comp}}} \sup_{(z,\mathcal{S},\boldsymbol{\theta})\in\mathcal{P}(\gamma)} \mathbb{E}[p\ell(\hat{z}_{\text{comp}}, z)] \\ &\geq \inf_{\hat{z}_{\text{comp}}} \mathbb{E}[p\ell(\hat{z}_{\text{comp}}, z^*_{\text{comp}})|(z^*_{\text{comp}}, \mathcal{S}^*_{\text{comp}}, \boldsymbol{\theta}^*_{\text{comp}})] \geq 1.\end{aligned}$$

The constructions of $(z^*_{\text{stats}}, \boldsymbol{\theta}^*_{\text{stats}})$ and $(z^*_{\text{comp}}, \mathcal{S}^*_{\text{comp}}, \boldsymbol{\theta}^*_{\text{comp}})$ are the most critical steps. With good constructions, the lower bound "≥ 1" can be verified by classical statistical conclusions (e.g. Neyman-Pearson Lemma) or prior work (e.g. HPC Conjecture).

A notable detail in the proof of statistical impossibility is the arbitrariness of

$\mathcal{S}^*$. The first infimum over $\mathcal{P}_{\mathcal{S}}(\gamma)$ in the minimax rate (3.10) requires that the lower bound (3.20) holds for any $\mathcal{S}^* \in \mathcal{P}_{\mathcal{S}}(\gamma)$. The arbitrary choice of $\mathcal{S}^*$ brings extra difficulties in the parameter construction, and consequently a non-trivial $\boldsymbol{\theta}^*_{\text{stats}} \neq \mathbf{1}$ is chosen to address the arbitrariness. Previous TBM construction in the proof of (Han et al., 2022a, Theorem 6) with $\boldsymbol{\theta}^*_{\text{stats}} = \mathbf{1}$ is no longer applicable in our case. Meanwhile, our construction $(z^*_{\text{comp}}, \mathcal{S}^*_{\text{comp}}, \boldsymbol{\theta}^*_{\text{comp}})$ leads to a rank-2 mean tensor to relate the HPC Conjecture while TBM (Han et al., 2022a, Theorem 7) constructs a rank-1 mean tensor. Hence, we emphasize that dTBM-specific techniques are required to obtain our impossibility results, though the proof idea is common for minimax lower bound analysis.

Proof sketch of Theorem 3.12

The proof of Theorem 3.12 is inspired by the proof idea of (Gao et al., 2018, Lemma 1). The extra difficulties are the angle gap characterization and multilinear algebra property in tensors; we address both challenges in our proof. Specifically, we control the misclustering error by the estimation error of $\hat{\mathcal{X}}$ calculated in Step 2 of Sub-algorithm 1. We prove the following inequality

$$\begin{aligned}
\ell(z^{(0)}, z) &\lesssim \frac{1}{p} \min_{\pi \in \Pi} \sum_{i: z^{(0)}(i) \neq \pi(z(i))} \theta(i)^2 \\
&\lesssim \frac{\sigma^2 r^{K-1}}{\Delta^2_{\min} p^K} \|\hat{\mathcal{X}} - \mathcal{X}\|_F^2 \\
&\lesssim \frac{r^K p^{-K/2}}{\text{SNR}},
\end{aligned} \tag{3.21}$$

where $\mathcal{X} = \mathbb{E}\mathcal{Y}$ is the true mean. The first inequality in (3.21) holds with the assumption $\min_{i \in [p]} \theta(i) \geq c > 0$ in Theorem 3.12. The second inequality relies on the key Lemma 3.4, which indicates

$$\min_{z(i) \neq z(j)} \|[\boldsymbol{X}_{i:}]^s - [\boldsymbol{X}_{j:}]^s\| \gtrsim \Delta_{\min}, \tag{3.22}$$

where $\boldsymbol{X} = \text{Mat}(\mathcal{X})$. The most challenging part in the proof of Theorem 3.12 lies in the derivation of inequality (3.22) (or the proof of Lemma 3.4), in which the proof of Gao et al. (2018) is no longer applicable due to different angle gap assumption in our dTBM. To address the angle gap notion, we develop the extra padding technique in Lemma B.5 and balance assumption (3.6). Last, we finish the proof of Theorem 3.12 by showing the third inequality of (3.21) using (Han et al., 2022a, Proposition 1).

Proof sketch of Theorem 3.17

The proof of Theorem 3.17 is inspired by the proof idea of (Han et al., 2022a, Theorem 2). We develop extra polar-coordinate based techniques with angle gap characterization to address the nuisance degree heterogeneity. Recall the intermediate quantity, misclustering loss, defined in (3.11)

$$\begin{aligned} L^{(t)} &:= L(z, z^{(t)}) \\ &= \frac{1}{p} \sum_{i \in [p]} \theta(i) \sum_{b \in [r]} \mathbb{1}\left\{ z^{(t)}(i) = b \right\} \left\| \left[\boldsymbol{S}_{z(i):} \right]^s - [\boldsymbol{S}_{b:}]^s \right\|^2. \end{aligned}$$

We show that $L^{(t)}$ provides an upper bound for the misclustering error of interest via the inequality $\ell^{(t)} \leq \frac{L^{(t)}}{\Delta_{\min}^2}$ in Lemma 3.11. Therefore, it suffices to control $L^{(t)}$. Further, we introduce the oracle estimators for core tensor under the true cluster assignment via

$$\tilde{\mathcal{S}} = \mathcal{Y} \times_1 \boldsymbol{W}^T \times_2 \cdots \times_K \boldsymbol{W}^T,$$

where $\boldsymbol{W} = \boldsymbol{M} \left(\text{diag}(\mathbf{1}_p^T \boldsymbol{M}) \right)^{-1}$ is the weighted true membership matrix. Let $\boldsymbol{V} = \boldsymbol{W}^{\otimes(K-1)}$ denote the Kronecker product of $(K-1)$ copies of $\boldsymbol{W}$ matrices, and we define the t-th iteration quantities $\boldsymbol{W}^{(t)}, \boldsymbol{V}^{(t)}$ corresponding to $\boldsymbol{M}^{(t)}$ (or

equivalently $z^{(t)}$). To evaluate $L^{(t+1)}$, we prove the bound

$$\begin{aligned}
&\mathbb{1}\left\{z^{(t+1)}(i)=b\right\}\\
&=\mathbb{1}\left\{\|[\boldsymbol{Y}_{i:}\boldsymbol{V}^{(t)}]^s-[\boldsymbol{S}^{(t)}_{b:}]^s\|^2\le\|[\boldsymbol{Y}_{i:}\boldsymbol{V}^{(t)}]^s-[\boldsymbol{S}^{(t)}_{z(i):}]^s\|^2\right\}\\
&\le A_{ib}+B_{ib},
\end{aligned}\tag{3.23}$$

where $\boldsymbol{Y}=\mathrm{Mat}(\mathcal{Y})$, $\boldsymbol{S}=\mathrm{Mat}(\mathcal{S})$, $\boldsymbol{S}^{(t)}=\mathrm{Mat}(\mathcal{S}^{(t)})$ and

$$\begin{aligned}
A_{ib}&=\mathbb{1}\left\{\left\langle\boldsymbol{E}_{i:}\boldsymbol{V},\left[\tilde{\boldsymbol{S}}_{z(i):}\right]^s-\left[\tilde{\boldsymbol{S}}_{b:}\right]^s\right\rangle\lesssim-\left\|\left[\boldsymbol{S}_{z(i):}\right]^s-[\boldsymbol{S}_{b:}]^s\right\|^2\right\},\\
B_{ib}&=\mathbb{1}\left\{\left\|\left[\boldsymbol{S}_{z(i):}\right]^s-[\boldsymbol{S}_{b:}]^s\right\|^2\lesssim F^{(t)}_{ib}+G^{(t)}_{ib}+H^{(t)}_{ib}\right\}.
\end{aligned}$$

The terms $F^{(t)}_{ib}, G^{(t)}_{ib}, H^{(t)}_{ib}$ are controlled by $z^{(t)}, \mathcal{S}^{(t)}$; see the detailed definitions in (B.43), (B.44), (B.45). Note that the event A_{ib} only involves the oracle estimator independent of t, while all the terms related to the t-th iteration are in B_{ib}. Thus, the inequality (3.23) decomposes the misclustering loss in the $(t+1)$-th iteration into the oracle loss and the loss in t-th iteration. This decomposition leads to the separation of statistical error and computational error in the final upper bound of Theorem 3.17.

Specifically, we prove the contraction inequality

$$\begin{aligned}
L^{(t+1)}&\le M\xi+\rho L^{(t)},\\
\text{with }\xi&=\frac{1}{p}\sum_{i\in[p]}\theta(i)\sum_{b\in[r]}A_{ib}\left\|\left[\boldsymbol{S}_{z(i):}\right]^s-[\boldsymbol{S}_{b:}]^s\right\|^2,
\end{aligned}\tag{3.24}$$

where M is a positive constant, $\rho\in(0,1)$ is the contraction parameter, and we call ξ the oracle loss. Controlling the probability of event B_{ib} and obtaining the $\rho L^{(t)}$ term in the right hand side of (3.24) are the most challenging parts in the proof of Theorem 3.17. Note that the true and estimated core tensors are involved via their normalized rows such as $\boldsymbol{S}^s_{a:}, \tilde{\boldsymbol{S}}^s_{a:}, [\boldsymbol{S}^{(t)}_{a:}]^s$. The Cartesian coordinate based analysis in Han et al. (2022a) is no longer applicable in our case. Instead, we use the polar-coordinate based analysis and the geometry property of trigonometric

functions to derive the high probability upper bounds for $F_{ib}^{(t)}, G_{ib}^{(t)}, H_{ib}^{(t)}$.

Further, by sub-Gaussian concentration, we prove the high probability upper bound for oracle loss

$$\xi \lesssim \text{SNR}^{-1} \exp\left(-\frac{p^{K-1}\text{SNR}}{r^{K-1}}\right). \tag{3.25}$$

Combining the decomposition (3.24) and the oracle bound (3.25), we finish the proof of Theorem 3.17.

The proof of MLE error shares the similar idea as Theorems 3.12-3.17. We first show a weaker polynomial rate for MLE and then improve the rate from polynomial to exponential through the iterations. The only difference is that the MLE remains the same over iterations due to its global optimality. See Appendix B, Section B.2 for the detailed proof.

4 MULTIWAY CLUSTERING – LARGE-SCALE SPEED-UP

This chapter presents the *multiway clustering* method for large-scale tensor data. Full work *Large-scale multiway clustering with seeded clustering* (Hu, 2025) is published in *The Second Conference on Parsimony and Learning* (*Proceedings Track*), *JMLR*.

4.1 Introduction

Applications of multiway clustering arise in a wide range of studies, including social relation graph clustering Banerjee et al. (2007), brain network community detection in neuroscience Al-Sharoa et al. (2018), and hypergraph analysis (Zhen and Wang, 2023). Many tensor-based approaches Chien et al. (2018); Wang and Zeng (2019); Ke et al. (2019); Han et al. (2022a); Hu and Wang (2023); Luo and Zhang (2022) have been successfully developed to capture the higher-order structures and solve the multiway clustering task with theoretical guarantees. Particularly, *Tensor Block Model* (TBM, Wang and Zeng (2019)) is one of the most popular statistical models for multiway clustering. TBM is considered as a higher-order generalization of matrix stochastic block model (SBM). The goal of TBM is to identify the underlying community structure in all directions (Output in Figure 4.1) from noisy tensor observations (Input in Figure 4.1). Previous works Wang and Zeng (2019); Han et al. (2022a) have investigated the theoretical properties and developed clustering algorithms with accuracy guarantees under TBM.

Despite the rapid development of tensor methods, one significant drawback of most tensor-based algorithms, including the TBM algorithm in Han et al. (2022a), is their computational inefficiency. Computational costs for one algorithm cover several aspects: (1) the storage size of inputs, (2) the *peak* memory cost to implement all intermediate calculations, and (3) the running time to obtain outputs. Due to multiple orders, the computational costs for tensor-based algorithms will inflate polynomially faster than one-dimensional and matrix-based algorithms. For example, the typical one-dimensional spectral clustering (Von Luxburg, 2007)

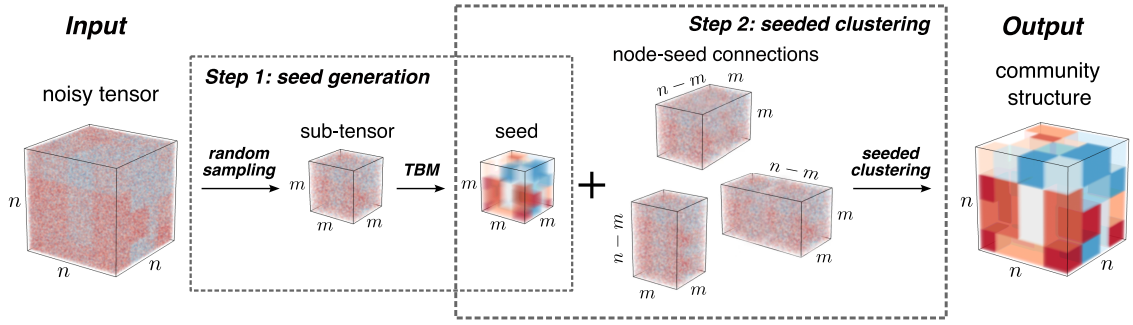

Figure 4.1: LS-TBM illustration for an order-3 tensor of with dimension n and seed size m on all modes.

for n entities has time complexity $\mathcal{O}(n^2)$ while the higher-order spectral clustering (Han et al., 2022a) for n entities on each of K directions has time complexity $\mathcal{O}(n^{K+1})$. This computational obstacle brings financial and time concerns to apply tensor-based algorithms to large-scale data in practice, even though tensor methods possess better theoretical properties.

In this work, we provide a large-scale multiway clustering framework, named *Large-Scale TBM* (*LS-TBM*), that substantially reduces the computational burden while maintaining similar accuracy as full TBM clustering. The key idea of LS-TBM is the utilization of *seeded clustering*: given a subset of node community assignments, called *seed*, we are able to infer the community assignments of remaining nodes. Specifically, we divide the multiway clustering into two steps: seed generation and seeded clustering. In Step 1, we randomly sample the node sets and apply the full TBM algorithm on the sub-tensor associated with subsampled nodes to obtain the seed. In Step 2, we infer the community assignments of all remaining nodes by comparing the node-seed connections with the block means estimated by the seed and sub-tensor. See Figure 4.1 for illustration.

Our contributions. We summarize the main contributions with our proposed LS-TBM below.

1. Our LS-TBM substantially reduces the algorithm time and space complexities

while maintaining the theoretical guarantees to exactly recover full community structures. Particularly, LS-TBM reduces polynomial complexities to logarithmic complexities with a strong signal-to-noise ratio (SNR) level. Table 4.1 shows the leading performance of LS-TBM in computation.

2. We establish accuracy guarantees for the two steps and the overall LS-TBM. We present the phase transition of LS-TBM performance with a key theoretical interplay between the SNR level and seed size m to the clustering accuracy. Figure 4.2 visualizes the phase transition and interplay.

3. We confirm the superior empirical performance of LS-TBM in large-scale numerical experiments.

Algorithm	Full TBM (Han et al., 2022a)	LS-TBM	LS-TBM (Strong SNR)	LS-TBM (Weak SNR)
Time Complexity	$\mathcal{O}(n^{K+1})$	$\mathcal{O}(m^{K+1})$	$\mathcal{O}(\log^{(K+1)/(K-1)} n)$	$\mathcal{O}(n^{(K+1)/2})$
Space Complexity	$\mathcal{O}(n^K)$	$\mathcal{O}(m^{K-1}n)$	$\mathcal{O}(n \log n)$	$\mathcal{O}(n^{(K-1)/2+1})$

Table 4.1: Time and space complexities of full TBM Han et al. (2022a) and LS-TBM on an order-K tensor with dimension n and seed size m. See Sections 4.2, 4.3, and Remark 4.8 for detailed discussions.

Related work. Here, we review several relevant lines of literature for comparison. The first line is about multiway clustering algorithms. One prominent approach is higher-order spectral clustering (Benson et al., 2015; Ghoshdastidar and Dukkipati, 2015; Ke et al., 2019; Wang and Li, 2020b; Han et al., 2022a; Luo and Zhang, 2022), which leverages low-dimensional spectral representations of data tensors. Higher-order spectral clustering extends traditional spectral clustering methods and incorporates various tensor decomposition methods such as CANDECOMP/PARAFAC (CP) decomposition (Hitchcock, 1927) and Tucker decomposition (Tucker, 1966). Specific strategies have also been developed for certain scenarios, including tensor-SCORE (Ke et al., 2019), angle-based iterations (Hu and Wang, 2023) for degree-corrected TBM, and the generalized linear model with alternating iterations (Wang

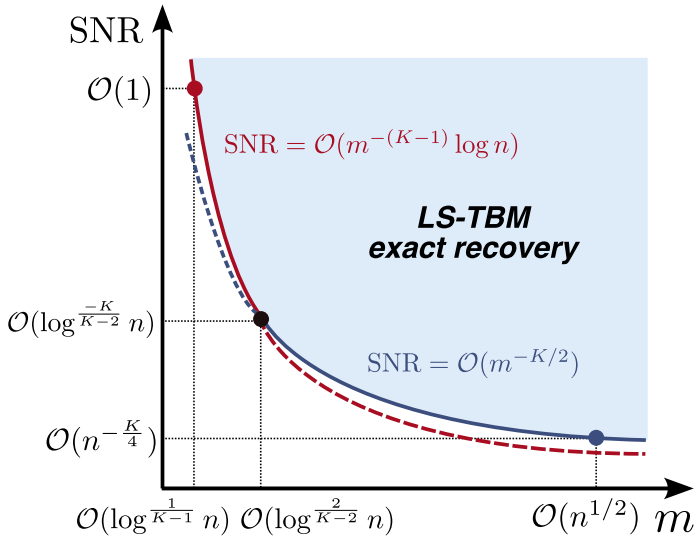

Figure 4.2: Phase transition of order-K LS-TBM performance with a theoretical interplay between the SNR level and seed size m. As SNR reduces from strong $\mathcal{O}(1)$ to weak $\mathcal{O}(n^{-K/4})$, LS-TBM requires an increasing seed size m from $\mathcal{O}(\log^{1/(K-1)} n)$ to $\mathcal{O}(n^{1/2})$ to the achieve exact recovery. When the combination (m, SNR) lies in the light blue area, LS-TBM fully recovers the community structure with a high probability.

and Li, 2020b) for binary tensor observations. Only a few works consider the computational difficulties in the applications with large-scale tensors while our work pays special attention to the computational feasibility, resource requirement, and running time in implementation. Through our framework, real applications with massive tensor data are able to practically enjoy the methodological achievements obtained by previous multiway clustering literature.

The second line focuses on scalable community detection. Several large-scale methods have been proposed for one-dimensional or matrix community detection problems, including fast pseudo-likelihood approaches (Amini et al., 2013; Wang et al., 2023), label propagation algorithms (Zhang et al., 2023; Traag and Šubelj, 2023), divide-and-conquer strategies (Mukherjee et al., 2021; Zhang et al., 2022b; Wu et al., 2023; Sayan Chakrabarty and Chen, 2023), and subsampling techniques (Zhang et al., 2022a; Deng et al., 2024). However, these methods rarely address the

multiway clustering problem for higher-order tensors. Given the unique algebraic properties of tensors, such extensions are non-trivial, and our work addresses this gap with a scalable tensor method.

The last line of research relates to the seed expansion. Generally, seeds refer to objects that require special attention in a system, and seed expansion propagates information from the seed to the entire system. In community detection, seed expansion reveals the full structure via node-seed connections through random walks (Su et al., 2017; Whang et al., 2016) or by comparing node characterizations (Whang et al., 2013; Xu et al., 2012). This concept has also been applied to other areas, including graph matching (Kazemi et al., 2015; Yu et al., 2021), essential node identification (Zhao et al., 2023), and relation extraction (Lang and Henderson, 2013). To the best of our knowledge, LS-TBM is the first work to adapt seed expansion to multiway clustering via seeded clustering.

Notation. We use lower-case letters (e.g., a) for scalars, lower-case bold letters (e.g., $\boldsymbol{a}$) for vectors, upper-case letters (e.g., S) for sets of integers, upper-case bold letters (e.g., $\boldsymbol{M}$) for matrices, and calligraphy letters (e.g., $\mathcal{X}$) for higher-order tensors with order equal or larger than 3. One exception is that we reserve letters 'c' (e.g., c, c_1) and 'C' (e.g., C, C_1) for generic small or large positive constants, respectively. For a positive integer n, let $[n] = \{1, \ldots, n\}$. For a set S, let $|S|$ denote the cardinality and $S(i)$ denote the i-th element in S. For two sequences $\{a_n, b_n\}$ of positive numbers, we denote $a_n \lesssim b_n$ or $a_n = \mathcal{O}(b_n)$ if $\lim_{n\to\infty} a_n/b_n \leq c$ and $a_n \gtrsim b_n$ or $a_n = \Omega(b_n)$ if $\lim_{n\to\infty} a_n/b_n \geq c$ for some constant $c > 0$; we denote $a_n = o(b_n)$ if $\lim_{n\to\infty} a_n/b_n = 0$; and we denote $a_n \asymp b_n$ if $a_n \gtrsim b_n$ and $a_n \lesssim b_n$. We use $\mathbb{1}\{\cdot\}$ for the 0-1 indicator function, $\|\cdot\|_2$ for the ℓ_2 norm, $\langle \cdot, \cdot \rangle$ for the inner product, $\otimes$ for the Kronecker product, $\mathbf{1}_n \in \mathbb{R}^n$ for the vector with all elements 1, $\boldsymbol{I}_n \in \mathbb{R}^{n\times n}$ for the identity matrix, and $\mathrm{diag}(\boldsymbol{v}) \in \mathbb{R}^{n\times n}$ for the diagonal matrix with elements in a vector $\boldsymbol{v} \in \mathbb{R}^n$. Consider an order-$K$ $(n_1, \ldots, n_K)$-dimensional tensor $\mathcal{X} \in \mathbb{R}^{n_1\times\cdots\times n_K}$ with entries $\mathcal{X}(i_1, \ldots, i_K)$ for $i_k \in [n_k], k \in [K]$. We use colon ': ' as a shorthand representing all entries in a particular direction. For example, $\mathcal{X}(i_1, \ldots, i_{K-1}, :) \in \mathbb{R}^{n_K}$ is a tensor fiber, and $\mathcal{X}(i_1, :, \ldots, :) \in \mathbb{R}^{n_2\times\cdots\times n_K}$ is a tensor

slice of the first mode. For a matrix $\boldsymbol{X}$, $\boldsymbol{X}_{i,:}$ and $\boldsymbol{X}_{:,j}$ are the i-th row and j-th column, respectively. We use $\text{Mat}_k(\cdot)$ for tensor unfolding operation along the k-th mode and $\text{Mat}_l(\mathcal{X}) \in \mathbb{R}^{n_l \times \prod_{k \neq l} n_k}$. Consider node subsets $S_k \in [n_k]$ for $k \in [K]$. We define sub-tensor associated with $\{S_k\}_{k\in[K]}$ as

$$\mathcal{X}(S_1, \ldots, S_k) = [\![\mathcal{X}(S_1(j_1), \ldots, S_K(j_K))]\!]_{j_k \in [|S_k|], k \in [K]} \in \mathbb{R}^{|S_1| \times \cdots \times |S_K|}.$$

The multilinear multiplication of a core tensor $\mathcal{C} \in \mathbb{R}^{r_1 \times \cdots \times r_K}$ by matrices $\boldsymbol{M}_k \in \mathbb{R}^{n_k \times r_k}$ results in an order-K $(n_1, \ldots, n_K)$-dimensional tensor $\mathcal{X}$, denoted as $\mathcal{X} = \mathcal{C} \times_1 \boldsymbol{M}_1 \times_2 \cdots \times_K \boldsymbol{M}_K$, with entries

$$\mathcal{X}(i_1, \ldots, i_K) = \sum_{j_1, \ldots, j_K} \mathcal{C}(j_1, \ldots, j_K) \boldsymbol{M}_1(i_1, j_1) \cdots \boldsymbol{M}_K(i_K, j_K).$$

Last, we drop the subscript $k \in [K]$ for any sequence with K elements $\{n_k\}_{k\in[K]}$ and use following shorthands for $\{n_k\}$: $n_* = \prod_{k\in[K]} n_k, n_{-l} = \prod_{k\in[K],k\neq l} n_k, \bar{n} = \max_{k\in[K]} n_k, \underline{n} = \min_{k\in[K]} n_k$.

4.2 Large-Scale Tensor Block Model Framework

Before the LS-TBM algorithm, we recall the TBM (Wang and Zeng, 2019) formula for setup. Consider an order-K tensor observation $\mathcal{Y} \in \mathbb{R}^{n_1 \times \cdots \times n_K}$. Suppose that there are r_k communities on the k-th mode of $\mathcal{Y}$. Let $z_k \colon [n_k] \mapsto [r_k], k \in [K]$ denote the community assignment functions. We say $\mathcal{Y}$ follows the TBM if

$$\mathcal{Y} = \mathcal{X} + \mathcal{E}, \quad \text{where} \quad \mathcal{X} = \mathcal{C} \times_1 \boldsymbol{M}_1 \times_2 \cdots \times_K \boldsymbol{M}_K, \tag{4.1}$$

$\mathcal{X} = \mathbb{E}[\mathcal{Y}] \in \mathbb{R}^{n_1 \times \cdots \times n_K}$ is the mean signal tensor, $\mathcal{E} \in \mathbb{R}^{n_1 \times \cdots \times n_K}$ is a noise tensor whose entries are independent and zero-mean with variance upper bounded by σ^2, $\mathcal{C} \in \mathbb{R}^{r_1 \times \cdots \times r_K}$ is the core tensor collecting the block means among communities, and $\boldsymbol{M}_k \in \{0,1\}^{n_k \times r_k}, k \in [K]$ are membership matrices such that $\boldsymbol{M}_k(i,a) = \mathbb{1}\{z_k(i) = a\}$. The community assignments $\{z_k\}$ or $\{\boldsymbol{M}_k\}$ are main parameters of interests.

We focus on high dimensional cases where n_k's increase while $r_k = \mathcal{O}(1)$.

The key idea of LS-TBM to alleviate computational burden is dividing the *high-dimensional and expensive* full algorithm into two *low-dimensional and cheap* steps: seed generation and seeded clustering. For a better narrative, we firstly introduce seeded clustering and then discuss seed generation.

Seeded clustering

Seeded clustering aims to extend the seed, i.e., the partial community assignments, to the full community structure. Taking the seed as "true labels", seeded clustering in fact transfers the unsupervised clustering problem to a relatively easy, supervised classification task.

Specifically, for illustration, we consider the seeded clustering on the first mode with a perfect seed $\{\tilde{z}_k \colon S_k \mapsto [r_k], S_k \subset [n_k]\}$ such that $\tilde{z}_k(i) = z_k(i)$ and $|S_k| = m_k$ for all $i \in S_k, k \in [K]$. Based on formula (4.1), the TBM structure still holds on the sub-tensor associated with $\{S_k\}$

$$\mathcal{X}(S_1, \ldots, S_K) = \mathcal{C} \times_1 \boldsymbol{M}_1(S_1, :\,) \times_2 \cdots \times_K \boldsymbol{M}_K(S_K, :\,),$$

where $\boldsymbol{M}_k(S_k, :\,) \in \{0,1\}^{m_k \times r_k}$ are membership matrices corresponding to $\{\tilde{z}_k\}$. By tensor multiplication, we have

$$\boldsymbol{C}_1 = \boldsymbol{W}_1^T \boldsymbol{X}_1(S_1, S_2 \times \cdots \times S_K)[\boldsymbol{W}_2 \otimes \cdots \otimes \boldsymbol{W}_K]^T, \tag{4.2}$$

where $\boldsymbol{C}_1 = \text{Mat}_1(\mathcal{C}) \in \mathbb{R}^{r_1 \times r_{-1}}$, $\boldsymbol{X}_1(S_1, S_2 \times \cdots \times S_K) = \text{Mat}_1(\mathcal{X}(S_1, \ldots, S_K)) \in \mathbb{R}^{m_1 \times m_{-1}}$ are mode-1 matricizations of the core tensor and the sub-tensor, respectively, and

$$\boldsymbol{W}_k = \boldsymbol{M}_k(S_k, :)(\text{diag}(\mathbf{1}_{m_k}^T \boldsymbol{M}_k(S_k, :)))^{-1}, \quad k \in [K],$$

are weighted membership matrices such that $\boldsymbol{W}_k^T \boldsymbol{M}_k(S_k, :) = \boldsymbol{I}_{r_k}$. The matrix $\boldsymbol{C}_1$ has r_1 unique rows, and the row $\boldsymbol{C}_1(a, :\,)$ collects the connections between the a-th community on the first mode to all r_{-1} communities on the other modes. Notice

that, for an arbitrary $j \in [n_1]/S_1$, we have

$$\boldsymbol{C}_1(z_1(j), :) = \boldsymbol{x}_j \coloneqq \boldsymbol{X}_1(j, S_2 \times \cdots \times S_K)[\boldsymbol{W}_2 \otimes \cdots \otimes \boldsymbol{W}_K]^T,$$

where $\boldsymbol{X}_1(j, S_2 \times \cdots \times S_K) = \text{Mat}_1(\mathcal{X}(j, S_2, \ldots, S_K)) \in \mathbb{R}^{r-1}$ collects the *node-seed connections* between the node j and seed node subsets $S_2, \ldots, S_K$. Then, we are able to obtain the community assignment for node j, $z_1(j)$, by comparing the aggregated node-seed connection vector $\boldsymbol{x}_j$ with r_1 rows of $\boldsymbol{C}_1$. Our seeded clustering is a supervised procedure. Instead of calculating the pairwise similarities among feature vectors $\boldsymbol{x}_j$'s, we classify the nodes via the similarities between $\boldsymbol{x}_j$'s and r_1 "true" reference centroids in $\boldsymbol{C}_1$. This supervised nature makes seeded clustering more computationally efficient than typical unsupervised clustering (e.g., k-means) who calculates similarities between all pairs of feature vectors. In practice, we use the noisy observation and imperfect seeds for estimation. See detailed procedures in the Sub-algorithm 2 of Algorithm 4.

Seed generation

Seed generation aims to obtain the partial community assignments $\{\tilde{z}_k \colon S_k \mapsto [r_k], S_k \subset [n_k]\}$. Seed provides a low-dimensional sketch of the whole community structure and will serve as the "true labels" for the following seeded clustering. As a pivot linking two steps, seed plays a critical role that determines the final clustering accuracy. A poor-quality seed misaligned with true z_k can mislead the assignments for remaining nodes. The misalignment in the seed will even be exaggerated in the full assignments due to the large dimension, making it more difficult to achieve overall exact recovery. A high-quality seed should satisfy three conditions: 1. $\{S_k\}$ are representative enough to cover all $\{r_k\}$ communities, 2. $\{\tilde{z}_k\}$ are accurate enough, and 3. sizes of the seed $\{m_k\}$ are proper.

Condition 1 is a basic requirement. It is impossible to obtain an accurate estimate of $\boldsymbol{C}_1(a, :)$ in (4.2) if none of a-th community members is in S_1. Fortunately, the *uniform sampling* addresses this requirement under most cases. If the true community structure is balanced, $\{S_k\}$ from uniform sampling are able to cover all communities

with a high probability. See Section 4.3 for the theoretical proof. Another advantage of uniform sampling is that no extra prior calculation is needed. Other sampling schemes (Rahmani et al., 2020; Xia and Yuan, 2021) are sophisticated but lead extra computation burdens. Condition 2 also ensures the estimation accuracy of $\boldsymbol{C}_1$ in (4.2). To achieve a high accuracy for seed assignments, we leverage the optimal full TBM algorithm Han et al. (2022a) on the low-dimensional sub-tensor. We recall the algorithm of Han et al. (2022a) as Algorithm 5 in Appendix. Condition 3 is the trickiest one and seed sizes $\{m_k\}$ are the key hyperparameters of the LS-TBM framework. If we choose large m_k's, the computational burden will increase due to the usage of full TBM on the sub-tensor. Whereas, small m_k's will lead the failure of Conditions 1 and 2 simultaneously. The choice of $\{m_k\}$ relates to the computational consideration, the signal level in $\mathcal{C}$, and the LS-TBM accuracy. For the question "how to choose proper $\{m_k\}$", we defer the discussion to Sections 4.3 and 4.4. Full procedures of LS-TBM framework are in Algorithm 4.

Remark 4.1 (Complexity of LS-TBM Algorithm 4). We analyze both time and space complexities, emphasizing the algorithm operation speed and memory cost, respectively. Consider the balanced case where $n_k \asymp n, m_k \asymp m$ for all $k \in [K]$ and assume $\{r_k\}, K$ as constants. The time complexity of Algorithm 4 is $\mathcal{O}(m^{K+1}+m^K+m^{K-1}n)$. The term $\mathcal{O}(m^{K+1})$ comes from the application of Algorithm 5 in Line 2. The other terms $\mathcal{O}(m^K + m^{K-1}n)$ come from the estimation of $\mathcal{C}$ (Line 3) and the aggregation (Line 5) in the seeded clustering step. The space complexity of Algorithm 4 is $\mathcal{O}(m^K + m^{K-1}n)$, dominated by the storage of the sub-tensor $\mathcal{Y}(S_1, \ldots, S_k)$ (Line 3) and node-seed connections $\mathcal{Y}(S_1, \ldots, S_k^c, \ldots, S_K)$ (Line 5). See comparison with full TBM in Table 4.1.

Algorithm 4 Large-scale multiway clustering under tensor block model

Sub-algorithm 1: Seed generation

Input: Observation $\mathcal{Y}$, number of communities $\{r_k\}$, sizes of the seed $\{m_k\}$, number of iterations T, relaxation factor in k-means: $M > 1$ (T, M are hyperparameters only for Algorithm 5).

1: Uniformly sample the node subsets S_k from $[n_k]$ such that $|S_k| = m_k$, for all modes $k \in [K]$.

2: Apply full TBM Algorithm 5 with $\{r_k\}, T, M$ on the sub-tensor $\mathcal{Y}(S_1, \ldots, S_K) \in \mathbb{R}^{m_1 \times \cdots \times m_K}$.

Output: Seed $\{\tilde{z}_k : S_k \mapsto [r_k]\}$ output by Algorithm 5.

Sub-algorithm 2: Seeded clustering

Input: Observation $\mathcal{Y}$, seed $\{\tilde{z}_k : S_k \mapsto [r_k]\}$ from Sub-algorithm 1.

3: Obtain the estimated core tensor with the seed $\{\tilde{z}_k\}$ as

$$\tilde{\mathcal{C}} = \mathcal{Y}(S_1, \ldots, S_K) \times_1 \tilde{\boldsymbol{W}}_1^T \times_2 \cdots \times_K \tilde{\boldsymbol{W}}_K^T, \text{ where } \tilde{\boldsymbol{W}}_k = \tilde{\boldsymbol{M}}_k(S_k, :)(\text{diag}(\boldsymbol{1}_{m_k}^T \tilde{\boldsymbol{M}}_k(S_k, :)))^{-1},$$

and $\tilde{\boldsymbol{M}}_k(S_k, :) \in \{0, 1\}^{m_k \times r_k}$ are membership matrices spanned by $\tilde{z}_k$ for $k \in [K]$.

4: **for** $k = 1$ to K **do**

5: Let $S_k^c = [n_k]/S_k$ be the complement of S_k. Calculate the aggregated tensor observation

$$\mathcal{A}_k = \mathcal{Y}(S_1, \ldots, S_{k-1}, S_k^c, S_{k+1}, \ldots, S_K) \times_1 \tilde{\boldsymbol{W}}_1^T \times_2 \cdots \times_{k-1} \tilde{\boldsymbol{W}}_{k-1}^T \times_{k+1} \tilde{\boldsymbol{W}}_{k+1}^T \times_{k+2} \cdots \times_K \tilde{\boldsymbol{W}}_K^T.$$

6: Calculate matricizations $\tilde{\boldsymbol{C}}_k = \text{Mat}_k(\tilde{\mathcal{C}}) \in \mathbb{R}^{r_k \times r_{-k}}$ and $\boldsymbol{A}_k = \text{Mat}_k(\mathcal{A}_k) \in \mathbb{R}^{(n_k - m_k) \times r_{-k}}$. The rows of $\tilde{\boldsymbol{C}}_k$ serve as r_k reference community centroids, and rows of $\boldsymbol{A}_k$ serve as aggregated feature vectors of remaining $n_k - m_k$ nodes waiting for classification.

7: **for** $j \in S_k^c$ **do**

8: Obtain the assignment for j as $\hat{z}_k(j) = \arg\min_{a \in [r_k]} \|\tilde{\boldsymbol{C}}_k(a, :) - \boldsymbol{A}_k(j, :)\|_2^2$.

9: **end for**

10: Combining seed assignments, obtain the full assignment $\hat{z}_k$ such that $\hat{z}_k(i) = \tilde{z}_k(i)$ for $i \in S_k$.

11: **end for**

Output: Estimated full community assignments $\{\hat{z}_k : [n_k] \mapsto [r_k]\}$.

4.3 Theoretical Guarantees

We start the theoretical analysis with several definitions and a general assumption. First, we define the clustering evaluation metrics. Take mode-1 as an example. The *misclassification error* of $\hat{z}_1$ is

$$\ell(\hat{z}_1, z_1) = \frac{1}{n_1} \min_{\pi \in \Pi} \sum_{i \in [n_1]} \mathbb{1}\{\hat{z}_1(i) \neq \pi \circ z_1(i)\},$$

where $\pi : [r_1] \mapsto [r_1]$ is a label permutation function, Π collects all possible permutations on $[r_1]$, and $\circ$ denotes the composition operation. We say the estimate $\hat{z}_1$ exactly recovers the true assignment z_1 if $\ell(\hat{z}_1, z_1) < 1/n_1$. Another metric is the *misclassification loss* of estimate $\hat{z}_1$, which is defined as

$$L(\hat{z}_1, z_1) = \frac{1}{n_1} \min_{\pi \in \Pi} \sum_{i \in [n_1]} \mathbb{1}\{\hat{z}_1(i) \neq \pi \circ z_1(i)\} \|\mathsf{Mat}_1(\mathcal{C})(\pi \circ z_1(i), :) - \mathsf{Mat}_1(\mathcal{C})(\hat{z}_1(i), :)\|_2^2,$$

where $\mathcal{C}$ is the core tensor in TBM (4.1). We drop the true assignment in misclassification error and loss, e.g., $\ell(\hat{z}_1)$ and $L(\hat{z}_1)$, for simplicity. Second, we define the *signal-to-noise ratio* (*SNR*) of TBM as

$$\text{SNR} = \frac{\Delta_{\min}^2}{\sigma^2}, \text{ where } \Delta_{\min}^2 = \min_{k \in [K]} \Delta_k^2, \text{ and } \Delta_k^2 = \min_{a \neq b \in [r_k]} \|\mathsf{Mat}_k(\mathcal{C})(a, :) - \mathsf{Mat}_k(\mathcal{C})(b, :)\|_2^2,$$

and σ^2 is the variance upper bound for the independent entries in the noise tensor $\mathcal{E}$ of (4.1).

Assumption 4.2 (Balanced communities)**.** *There exist universal positive constants* α_1, α_2 *such that* $\alpha_1 n_k / r_k \leq \sum_{i \in [n_k]} \mathbb{1}\{z_k(i) = a\} \leq \alpha_2 n_k / r_k$*, for all* $a \in [r_k], k \in [K]$.

Assumption 4.2 is mild. Such assumption is common in multiway clustering literature Wang and Zeng (2019); Ke et al. (2019); Han et al. (2022a); Hu and Wang (2023). Practically, we ignore this balance assumption in the numerical experiments; i.e., we randomly generate synthetic community assignments without size constrains. See Section 4.4 for details.

Accuracy of seeded clustering

The accuracy of seeded clustering relies on the seed. Next assumption requires the seed balance.

Assumption 4.3 (Balanced communities in seed node subsets)**.** *There exist universal positive constants α_3, α_4 such that $\alpha_3 m_k/r_k \leq \sum_{i\in S_k} \mathbb{1}\{z_k(i) = a\} \leq \alpha_4 m_k/r_k$, for all $a \in [r_k], k \in [K]$.*

Let $z_k^c \colon [n_k]/S_k \mapsto [r_k]$ denote the community assignments for remaining nodes, for all $k \in [K]$. Now, we are ready to present the accuracy of seeded clustering.

Theorem 4.4 (Accuracy of seeded clustering)**.** *Suppose that the tensor observation $\mathcal{Y}$ follows TBM* (4.1) *with number of communities $\{r_k\}$. Let $\{\tilde{z}_k : S_k \mapsto [r_k]\}$ be the given seed, and let $\{\hat{z}_k : [n_k] \mapsto [r_k]\}$ denote the output of Sub-algorithm 2 in Algorithm 4 with $\mathcal{Y}$ and $\{\tilde{z}_k\}$. Suppose that Assumptions 4.2 and 4.3 hold. If SNR and the seed satisfy following conditions with some positive large constant C and small constant c:*

$$SNR \geq C\frac{\bar{m}}{m_*}\log \bar{m}, \quad L(\tilde{z}_k) \leq c\frac{\Delta_{\min}^2}{\sqrt{\bar{m}}}, \quad \text{for } k \in [K], \tag{4.3}$$

we have

$$\ell(\hat{z}_k^c) \leq C' SNR^{-1} \exp\left(-c' m_{-k} SNR\right), \quad \text{for } k \in [K],$$

with probability at least $1 - \exp(-c''\underline{m}) - \exp(-c'''m_{-k}\Delta_{\min}^2/\sigma^2)$ and C', c', c'', c''' are positive constants.

Consider the balanced case where $n_k \asymp n$ and $m_k \asymp m$ for all $k \in [K]$. The first condition in (4.3) requires the SNR level to be at least $\mathcal{O}(m^{-(K-1)}\log m)$. This signal requirement aligns with the SNR requirement in the guarantee of the higher-order Lloyd algorithm (Han et al., 2022a) with a tensor of dimension m. Such condition indicates that our seeded clustering shares a similar theoretical performance as the Lloyd iteration. The second condition in (4.3) provides the boundary of the seed's accuracy to guarantee a good performance in seeded clustering. Note that we do not assume the seed generation approach for Theorem 4.4. As long as the

seed satisfies the condition (4.3) and other assumptions hold, the seeded clustering achieves a fast exponential error rate $\mathcal{O}(\exp(-m^{K-1}))$ given a fixed SNR.

Quality of seed generation

As described in Section 4.2 and indicated by Theorem 4.4, we consider the seed quality from two aspects: the coverage of community structure and the accuracy of seed assignments.

Theorem 4.5 (Balanced communities in uniformly sampled node subsets)**.** *Consider the node subsets $\{S_k\}$ obtained by uniform sampling from $\{[n_k]\}$, such that $|S_k| = m_k$ for all $k \in [K]$. Suppose that Assumption 4.2 holds. With probability at least $1-C\exp(-cm_k)$, we have*

$$\frac{\alpha_3 m_k}{r_k} \leq \sum_{i \in S_k} \mathbb{1}\{z_k(i) = a\} \leq \frac{\alpha_4 m_k}{r_k}, \quad \textit{for all } a \in [r_k], k \in [K]$$

where C, c, α_3, α_4 are some universal positive constants.

By Theorem 4.5, we formally show that the simple uniform sampling selects node subsets with balanced true community structures. That is, node subsets $\{S_k\}$ generated by the Line 1 in Algorithm 4 satisfies Assumption 4.3, which is required by the seeded clustering guarantee Theorem 4.4.

Next, for accuracy, we present the theoretical guarantee in Han et al. (2022a) as a corollary under our context.

Corollary 4.6 (Accuracy of seed assignments Han et al. (2022a))**.** *Suppose that the tensor observation $\mathcal{Y}$ follows TBM (4.1) with number of communities $\{r_k\}$. Let $\{\tilde{z}_k : S_k \mapsto [r_k]\}$ denote the output of Sub-algorithm 1 in Algorithm 4 with given inputs $\mathcal{Y}$, $\{m_k\}$, T, and M. Suppose that Assumption 4.2 holds. If SNR and number of iterations T satisfy the following conditions with some positive large constants C, C':*

$$\textit{SNR} \geq C m_*^{-1/2}, \quad T \geq C' \log \bar{m}, \tag{4.4}$$

we have

$$L(\tilde{z}_k) \le C''\sigma^2 \exp\left(-\frac{c_1 m_*}{\bar{m}} SNR\right) + c_2 \frac{\Delta_{\min}^2}{2^T} \le c\Delta_{\min}^2/\sqrt{\bar{m}}, \quad \textit{for all } k \in [K], \quad (4.5)$$

with a probability at least $1 - \exp(-c'\underline{m}) - \exp(-c''m_{-k}\Delta_{\min}^2/\sigma^2)$. *Here,* c, c', c'', c'', c_i *are some small positive constants related to* M *and* C'' *is a large positive constant.*

Under the balanced case, the first condition in (4.4) requires a stronger SNR level at rate $\mathcal{O}(m^{-K/2})$, compared with Theorem 4.4. Such stronger SNR requirement is due to the spectral initialization in full TBM Algorithm 5. The second condition in (4.4) requires a large number of iterations for the iterative Lloyd step in Algorithm 5. Note that the upper bound (4.5) aligns with condition (4.3) in Theorem 4.4. Combining with Theorem 4.5, we show that our seed generation, Sub-algorithm 1 in Algorithm 4, is good enough to provide high-quality seeds for seeded clustering.

Overall accuracy of large-scale multiway clustering

Now, we show the overall accuracy of Algorithm 4, combining the theoretical guarantees for seeded clustering and seed generation.

Corollary 4.7 (Overall accuracy of LS-TBM)**.** *Suppose that the tensor observation* $\mathcal{Y}$ *follows TBM* (4.1) *with number of communities* $\{r_k\}$. *Let* $\{\hat{z}_k\}$ *denote the output of Algorithm 4 with given inputs* $\mathcal{Y}$, $\{m_k\}$, T, *and* M. *Suppose that Assumption 4.2 holds. If SNR and number of iterations* T *satisfy the following conditions with some positive large constants* C_1, C_2, C_3*:*

$$SNR \ge C_1 m_*^{-1/2} \vee C_2 \frac{\bar{m}}{m_*} \log \bar{m}, \quad T \ge C_3 \log \bar{m}, \quad (4.6)$$

we have

$$\ell(\hat{z}_k) \le C_4 SNR^{-1} \exp\left(-\frac{c_1 m_*}{\bar{m}} SNR\right) + c_2 \frac{m_k}{2^T n_k}, \quad \textit{for all } k \in [K] \quad (4.7)$$

with high probability at least $1 - C_5 \exp(-c_3 \underline{m}) - \exp(-c_4 \frac{m_*}{\bar{m}} SNR)$ *as* $m_k \to \infty, n_k \to \infty$ *and some positive constants* c_i*'s and* C_i*'s.*

Corollary 4.7 combines the results in seeded clustering and seed generation. Conditions (4.6) for SNR and number of iterations are inherited from the seed generation step, Corollary 4.6. The error rate in (4.7) can be obtained with Theorem 4.4 and Corollary 4.6 by the fact that $\ell(\hat{z}_k) = \frac{m_k}{n_k}\ell(\tilde{z}_k) + \frac{n_k - m_k}{n_k}\ell(\hat{z}_k^c)$.

Remark 4.8 (Interplay between seed size m and SNR to the exact recovery of LS-TBM). Corollary 4.7 indicates the mutual effects of the seed size m and SNR level to the misclassification error bound (4.7). Under the balanced case $n_k \asymp n, m_k \asymp m$, we investigate the interplay between m and SNR to the exact recovery of LS-TBM to answer the question "*how to select a proper* m" raised in Section 4.3. Here, we assume $K \geq 3$ and T is large enough to ignore the second term in the bound (4.7).

To achieve exact recovery with LS-TBM, i.e., $\ell(\hat{z}_k) \leq 1/n$ for all $k \in [K]$, we need following conditions to be fulfilled simultaneously by Corollary 4.7:

$$\text{SNR} \gtrsim m^{-K/2}, \quad \text{and} \quad m^{K-1}\text{SNR} \gtrsim \log n, \tag{4.8}$$

where the first inequality comes from (4.6) and the second inequality comes from (4.7). Based on (4.8), when $m \gtrsim \log^{2/(K-2)} n$, we need SNR $\gtrsim m^{-K/2}$ to achieve exact recovery; when $m \lesssim \log^{2/(K-2)} n$, exact recovery requires SNR $\gtrsim m^{-(K-1)} \log n$. Inequalities (4.8) also indicate the trade-off between m and SNR: a larger SNR level is needed if we want to use a smaller m in LS-TBM for multi-way clustering, and the TBM with a smaller SNR requires a larger seed size m for LS-TBM to recover. Figure 4.2 visualizes such phase transition of LS-TBM and the interplay (4.8). Besides the intersection point $\mathcal{O}(\log^{2/(K-2)} n)$, we also discuss two representative cases with strong and week SNRs.

First, consider the strong signal case with constant SNR $\asymp 1$. By the interplay (4.8), we only need a logarithmic seed size $m = \mathcal{O}(\log^{1/(K-1)} n)$ to achieve exact recovery. Then, based on Remark 4.1 and Han et al. (2022a), the time complexity to fully recover the community structure with full TBM, $\mathcal{O}(n^{K+1})$, dramatically drops to $\mathcal{O}(\log^{(K+1)/(K-1)} n)$ with LS-TBM. Also, the space complexity drops from $\mathcal{O}(n^K)$

to $\mathcal{O}(n \log n)$. This complexity comparison proves that LS-TBM has a huge potential to release the computational burden of multiway clustering while keeping the exact recovery performance.

Second, we consider a relatively weak signal case with SNR $\asymp n^{-K/4}$. We need a polynomial seed size $m = \mathcal{O}(n^{1/2})$ for exact recovery by (4.8). Under this case, time complexity drops from $\mathcal{O}(n^{K+1})$ to $\mathcal{O}(n^{K+1}/2)$ with LS-TBM, and space complexity reduces from $\mathcal{O}(n^K)$ and $\mathcal{O}(n^{(K-1)/2+1})$. We conclude that LS-TBM polynomially relieves the computational burden even under a weak SNR.

Table 4.1 summarizes the complexity comparison under both strong and weak SNR cases.

4.4 Numerical Experiments

Simulations

For simulations, we consider order-3 tensor observations from TBM (4.1) with $n_k = n, r_k = r$ and seed size $m_k = m$ for $k = 1, 2, 3$. We generate true assignments $\{z_k\}$ via random sampling from $[r]$ and set noise level $\sigma^2 = 1$. We use the clustering error rate (CER), i.e., one minus the adjusted Rand index, to measure clustering accuracy. The CER is equal to the misclassification error $\ell(\hat{z})$ up to a constant factor (Meilă, 2012). We report average statistics and standard deviations across 30 replications [1].

Our first experiment verifies the theoretical interplay in Section 4.3. We choose three SNR levels from strong to weak: $\mathcal{O}(1), \mathcal{O}(\log^{-3} n)$, and $\mathcal{O}(n^{-3/4})$. Based on Remark 4.8, the seed size thresholds are $m = \mathcal{O}(\log^{1/2} n), \mathcal{O}(\log^2 n)$, and $\mathcal{O}(n^{1/2})$, respectively. LS-TBM performance is expected to change rapidly around the thresholds. Figure 4.3 confirms such phase transition. Under the strong SNR $= \mathcal{O}(1)$ case, all LS-TBM algorithms converge except that with constant seed size $m = \mathcal{O}(1)$; under the weak SNR $= \mathcal{O}(n^{3/4})$, a significant performance gap is observed between $m = \mathcal{O}(\log^2 n)$ and $\mathcal{O}(n^{1/2})$. Additional experiments in Appendix further support

[1] LS-TBM software is available at `https://github.com/Marchhu36/LS-TBM`.

the rapid LS-TBM performance changes around the seed size threshold $m = \mathcal{O}(n^{1/2})$ under the weak SNR case.

Our second experiment compares the empirical performance of LS-TBM and full-TBM from accuracy (in CER), memory cost (peak RAM, in Mb), and running time (in second). We consider the strong SNR $= \mathcal{O}(1)$ cases. Experiments that exceed the memory limit 25 Gb will be accounted as infeasible. Figure 4.4 shows the leading advantage of LS-TBM in saving computational resources without accuracy sacrifices. The memory cost and running time of full TBM inflate at a polynomial rate, while the costs of LS-TBMs increase in an extremely slow rate. Such phenomenon agrees with the complexity comparison in Table 4.1. Moreover, the gap in computational cost between LS-TBM and full TBM increases significantly as the order K increases. See Appendix for the additional experiments with $K = 4, 5$. Therefore, we conclude that LS-TBM is favorable than full TBM for large-scale multiway clustering.

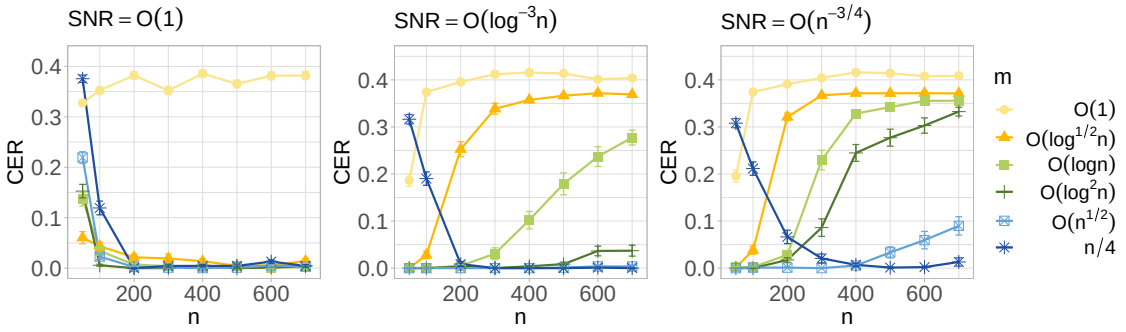

Figure 4.3: Interplay between the SNR and seed size m in LS-TBM. Set dimension n from 50 to 700, SNR from strong $\mathcal{O}(1)$ to weak $\mathcal{O}(n^{-3/4})$, and seed size m from $\mathcal{O}(1)$ to $n/4$.

Real data analysis

We apply our LS-TBM to the Uber Pickup Data Smith et al. (2017) in New York City (NYC) from April 2014 to August 2014. We organize the data as an order-3 tensor $\mathcal{Y} \in \mathbb{R}^{4392\times 445\times 559}$ with count entries recording the hourly amount of Uber pickups,

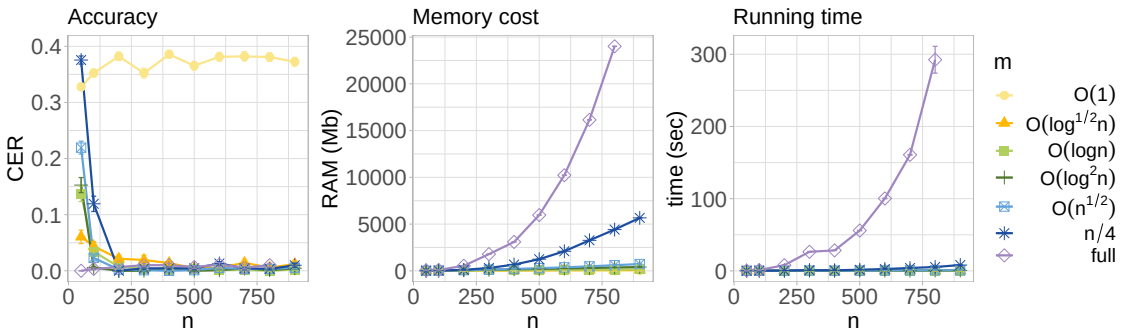

Figure 4.4: Empirical comparison between LS-TBMs and full TBM from accuracy (CER), memory cost (peak RAM, in Mb), and running time (in second). Set dimension n from 50 to 900, SNR $= \mathcal{O}(1)$, and seed size m from $\mathcal{O}(1)$ to $n/4$. Full TBM exceeds the memory upper limit of 25 Gb when $n > 800$.

where mode-1 refers to timestamps of 183 days $\times$ 24 hours, mode-2 and mode-3 refer to 445 latitude and 559 longitude coordinates, respectively. To mimic the application on personal laptop, we choose $m_1 = 0.1n_1, (m_2, m_3) = 0.5(n_2, n_3)$ for LS-TBM, taking peak RAM around 5 Gb and running time 25 seconds. Due to the large dimension, it is time-consuming to perform r selection on the entire $\mathcal{Y}$. We select $r_1 = 6$ based on the time combinations (Workday, Weekend) $\times$ (Early Morning, Daytime, Evening) and select $r_2 = r_3 = 4$ as the maximal numbers not leading singletons in LS-TBM clustering. For comparison, we still conduct the expensive full TBM with peak RAM around 27 Gb and running time 40 minutes.

Figure 4.5A shows the similar geographic community structure of NYC obtained by TBM and LS-TBM. The averaged CER between LS-TBM and full TBM on the last two modes (latitude and longitude) is 0.015. Particularly, both LS-TBM and full TBM identify an unique cluster highlighting the Manhattan district. For the first mode (timestamp), LS-TBM learns different time clusters compared to full TBM but captures major patterns. See detailed time cluster comparison in Appendix. Figure 4.5B shows three representative common time clusters: Cluster 1 (Early Morning $\times$ Weekdays), Cluster 2 (Daytime $\times$ Workdays), and Cluster 5 (Evening $\times$ Later Workdays). Meanwhile, by Figure 4.5C, the averaged Uber pickup amount in

Manhattan neighborhoods increases from early morning to evening. Overall, given the similar clustering performance but the huge gap in computational requirements, we say LS-TBM is more practically useful than full TBM for real-world large-scale multiway clustering tasks.

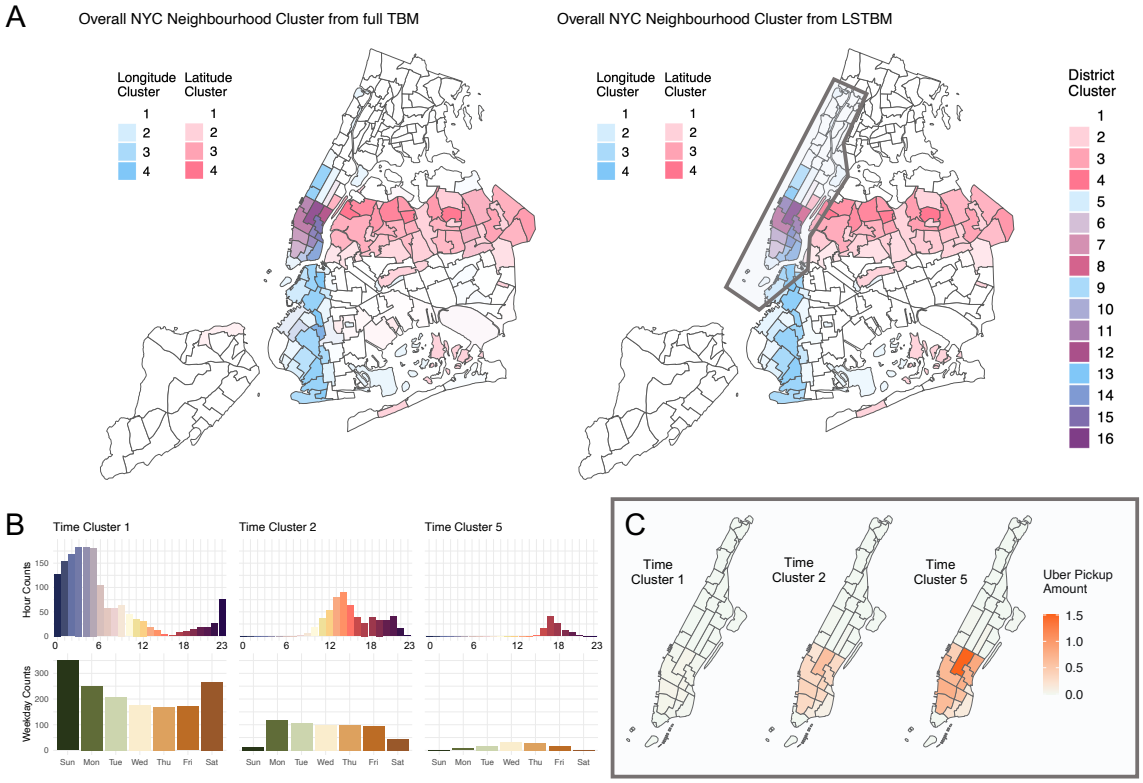

Figure 4.5: Multiway clustering on Uber Pickup Tensor. **A.** Geographic communities obtained by full TBM and LS-TBM. NYC is partitioned in 16 district clusters combined from 4 longitude clusters $\times$ 4 latitude clusters. **B.** Counts of 24 hours and 7 weekdays for timestamps in Time Clusters 1, 2, and 5 from LS-TBM. **C.** Estimated hourly Uber pickup amount in Manhattan neighborhoods under different time clusters.

4.5 Conclusion

This paper introduces an efficient multiway clustering framework, LS-TBM, for large-scale tensors. LS-TBM decomposes the high-dimensional and expensive TBM algorithm into two low-dimensional and cheap steps, seed generation and seeded clustering. Complexity analysis and accuracy guarantees theoretically proves the efficacy of LS-TBM: substantially reduces the computational cost while achieving a desirable clustering performance. In particular, under the strong SNR case, LS-TBM is able to exactly recover the community structure with logarithmic complexity, in contrast to the polynomial complexity required by the full TBM algorithm. Simulations and real data analysis using Uber Pickup data further empirically validate the superior efficiency of LS-TBM in terms of runtime and memory usage while maintaining a comparable accuracy to the full TBM algorithm.

LS-TBM is a generalizable framework with potential extensions in multiple directions. In the seed generation step, when additional prior community information is available, structural sampling may be preferable to uniform sampling for generating high-quality seeds. Furthermore, since the performance of LS-TBM only depends on seed quality, more sophisticated multiway clustering methods can be incorporated for seed generation in specific scenarios, such as cases with high sparsity and heavy-tailed observations. In the seeded clustering step, alternative distance metrics may be employed in Step 8 of Sub-algorithm 2 for different purposes. For instance, an angle-based comparison can be adopted for seeded clustering under the degree-corrected model, while people may use absolute distance for robust clustering. Therefore, we believe that LS-TBM serves as a strong foundation to the development of generalized multiway clustering methods for large-scale tensors.

5 TENSOR-BASED NETWORK QTL MAPPING

This chapter presents the *spectral network quantitative trait loci analysis* (snQTL) to map genetic loci affecting the gene co-expression network. Full work *A spectral framework to map QTLs affecting joint differential networks of gene co-expression* (Hu et al., 2024) is published in *PLoS Computational Biology*.

5.1 Introduction

The identification of genetic variants underlying complex phenotypic traits has been a pivotal area in genetics research for decades. Genome-wide association studies (GWASs) have identified important genetic variants by detecting statistical association between phenotypes and genotypes in outbred populations Visscher et al. (2012). Likewise, quantitative trait locus (QTL) mapping in experimentally crossbred organisms allows researchers to shuffle genetic backgrounds meiotically and test for associations between measurable phenotypes and chromosomal regions. However, both GWAS and QTL mapping are limited by the challenge of elucidating the mechanisms behind these genotype-phenotype associations, and the lack of sufficient functional information for many loci Zhu et al. (2016); Yuan et al. (2022a). Gene expression studies can bridge this gap between genotype and phenotype. To this end, expression quantitative trait locus (eQTL) analysis was developed to identify associations between genetic variants and gene expression levels Nica and Dermitzakis (2013). The eQTL studies have deepened our understanding of genotype $\rightarrow$ expression $\rightarrow$ phenotype mechanisms (Majewski and Pastinen, 2011; Peters et al., 2016; Zhu et al., 2016; Zhernakova et al., 2017). Existing eQTL methods have identified numerous genetic loci, categorized as cis- or trans-eQTL, that influence gene expression. The cis-eQTLs are located near the expressed gene on the same chromosome, and they typically directly affect the binding of transcription factors or chromatin proteins to DNA Ding et al. (2010); Guo et al. (2018). Growing evidence has indicated that cis-eQTLs do not fulfill the expected

role of linking genetic variants to target genes or pathways (Consortium, 2020; Yao et al., 2020; Hoskins et al., 2021; Dutta et al., 2022; Wang and Goldstein, 2020). Conversely, trans-eQTLs are typically located far from the expressed gene on the same chromosome or on different chromosomes. In most eQTL mapping studies, large blocks of linkage disequilibrium (LD) make it challenging to distinguish between a true cis-eQTL and a trans-eQTL that regulates a nearby gene on the same chromosome.

A key limitation of current eQTL studies is their focus on individual genes, but not on the network structure of gene co-expression. Gene co-expression networks are often represented by correlation matrices at the whole-transcriptome scale Stuart et al. (2003); Ruprecht et al. (2017); Gaiteri et al. (2014). Correlation among gene expressions may arise, for example, when multiple genes are co-regulated by the same transcription factor or participate in sequential regulatory cascades. Correlated expression can also arise from genetic linkage between separate regulatory cascades and from shared environmental effects, although in this case correlated expression do not necessarily imply direct functional interactions.

There is accumulating evidence that gene co-expression networks can differ between species Ovens et al. (2021); Crow et al. (2022) or populations Ovens et al. (2021), even in controlled environments. These differences suggest the gene co-expression network is evolvable, and hence most likely has a genetic basis. The co-expression might evolve with genetic variants, for instance, if transcription co-factor A modifies the effect of transcription factor B on target gene C's expression, then allelic differences at gene A can modify the correlation between expression of genes B and C. Mutations that alter gene linkage patterns (e.g., inversions or translocations) could also alter gene co-expression networks. This concept is similar to mapping epistatic eQTLs Kang et al. (2015), except that those studies (excluding work on highly prolific laboratory models) only rarely have the power to identify more than a few interacting genes. If genetic variants broadly alter co-expression network structure, eQTL or GWAS methods could in principle map these genetic loci. Analyzing these associations between genetic loci and co-expression network can reveal the network-level impact of quantitative trait loci, leading to new insights

into the genetic basis of complex traits. Developing efficient methods for network-based eQTL is a topic of great interest.

Recent studies have extended the concept of eQTL to co-QTLs Van Der Wijst et al. (2018); Baker et al. (2019); Yuan et al. (2022a); Li et al. (2023). These methods aim to identify genetic loci that explain coordinated changes in expression between pairs of genes. However, current co-QTL methods have several limitations. One of the challenges is the massive number of statistical tests needed, which increases quadratically with the number of gene pairs analyzed. Some methods restrict co-QTL searches to previously identified eQTLs Van Der Wijst et al. (2018); Li et al. (2023), while others prioritize gene pairs based on prior knowledge Yuan et al. (2022a). These approaches reduce testing burdens but may miss important co-QTLs. Furthermore, current co-QTL methods are limited by assuming linear models with additive effects Van Der Wijst et al. (2018); Baker et al. (2019); Yuan et al. (2022a); Li et al. (2023). The additive assumption neglects dominance, recessiveness, or even transgressive inheritance, hindering the ability to capture the full genetic influence on co-expression networks. Recent work on activity QTL (aQTL) (Hoskins et al., 2021) has extended traditional eQTL analysis by using gene activity scores inferred from co-expression networks. However, this method fails to accommodate multiple networks or relax the linear assumptions in co-QTL analysis. More powerful network-based eQTL methods are needed to address these issues.

In this paper, we propose a novel method called spectral network QTL (snQTL) to address these challenges. Our snQTL approach identifies the association between genotype and the entire co-expression network structure. The identified loci, which we also refer to as snQTLs, explain a fraction of the genetic variance of the entire gene co-expression network. The snQTL represent genetic variants that alter the global pattern of a network, while traditional co-QTLs represent genetic variants that alter the expression for only a particular pair of genes. Statistically, the key idea of the snQTL method is to use tensor spectral statistics to represent the joint difference in gene co-expression networks at each of many different loci. This approach reduces the number of tests to the number of genetic markers throughout

the genome of a recombinant hybrid population (used for mapping), and we allow for the simultaneous consideration of all active genes in the network. We also propose a permutation-based approach to obtain valid testing results that are robust to the data distribution. In addition to identifying snQTLs, our snQTL framework also outputs the joint differential networks, which represents the specific network patterns that are altered by genetic variants at the detected snQTLs. Our approach has the potential to be extended to mapping genetic effects on the architecture of microbiome co-occurrence networks and proteomic networks. We demonstrate the effectiveness of our method in the immune tissue gene expression data from a large genetic cross of three-spined stickleback fish (*Gasterosteus aculeatus*).

5.2 Results

Spectral network QTL framework

Fig 5.1 illustrates the main framework of our snQTL method. We take as input (i) expression read counts of p genes and (ii) genotypes of m genetic markers, from the same set of n individuals. The snQTL method then outputs two key results at each marker: (i) a p-value indicating the association significance between the co-expression network and the marker, and (ii) a joint differential network with nodes representing genes and edges representing associated effects.

The snQTL consists of three steps. First, we construct gene co-expression networks; see Step 0 in Fig 5.1. At each of the m markers, we group the samples by genotype (AA, AB, BB). For each group, we calculate a Pearson correlation matrix using only the gene expression data within that group. Let N_A, N_B, N_H denote the (unknown) population correlation matrices, where A and B denote homozygous genotypes and H denotes heterozygous genotype. We exclude within-chromosome correlations by setting the (j, k)-th entries in N_A, N_B, N_H to zero, if genes j and k are located on the same chromosome. The purpose of this 'set-to-zero' step is to focus on trans-snQTLs that affect between-chromosome co-expressions which are less likely to result from linkage disequilibrium (LD) and more likely to indicate

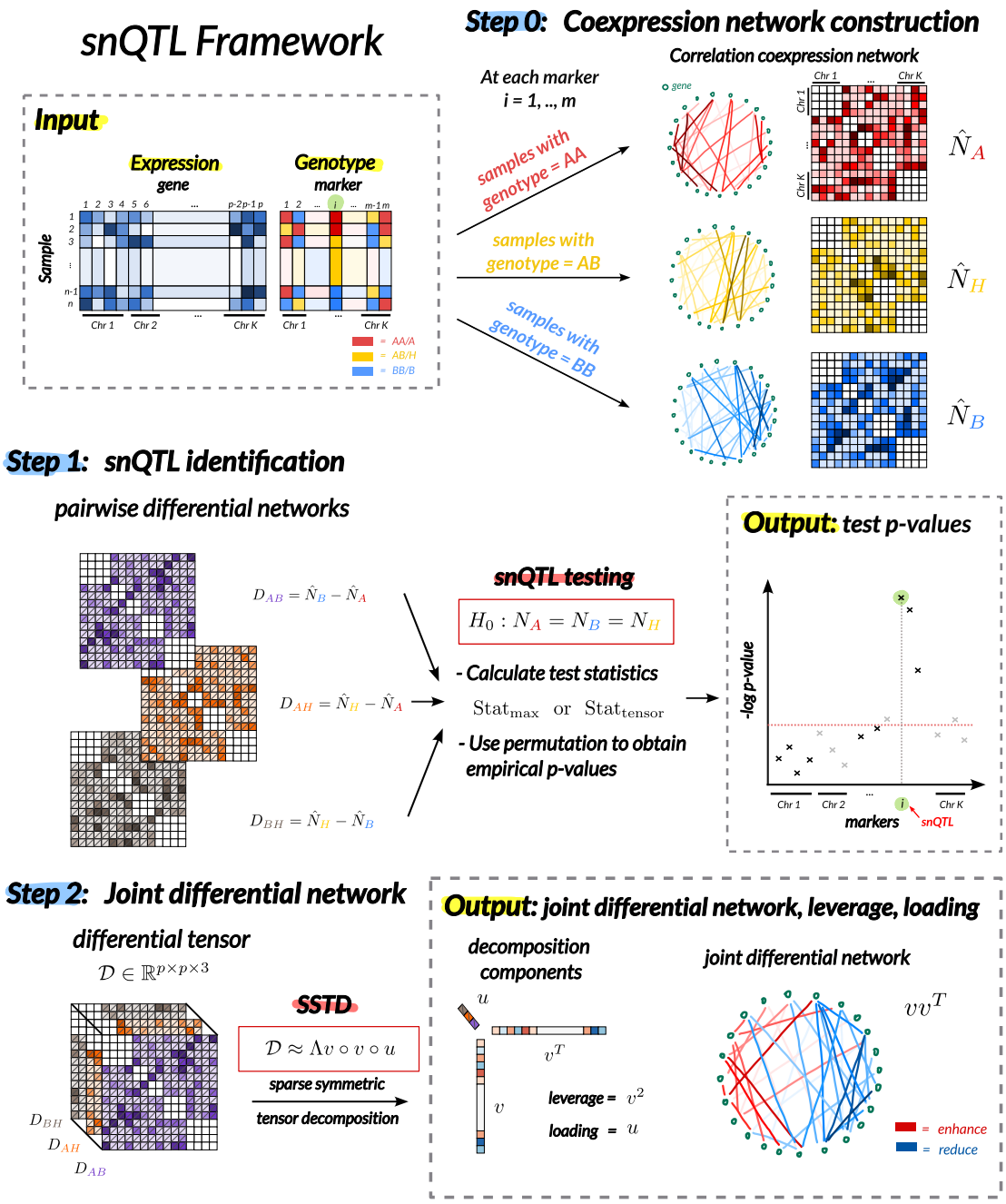

Figure 5.1: **The main idea of our snQTL framework.** Our snQTL framework takes as input (i) gene expression read counts and (ii) genotypes of genetic markers from the same set of samples. The snQTL approach consists of three steps: (0) co-expression network construction, (1) snQTL identification via hypothesis testing using multilinear spectral statistics, and (2) joint differential network estimation at associated loci via sparse symmetric tensor decomposition. At each marker, the output includes (i) a p-value indicating the association significance between the co-expression network and the marker, and (ii) a joint differential network with nodes representing genes and edges representing associated effects.

functional connections. In eQTL studies, trans-eQTLs often influence the expression or structure of transcription factors, ultimately affecting their ability to regulate distant genes Swanson-Wagner et al. (2009); Westra et al. (2013); Võsa et al. (2021). Although, in theory, a trans-eQTL could act on a nearby gene (e.g., if a transcription factor regulates a gene immediately adjacent on a chromosome), QTL mapping often cannot distinguish cis-eQTLs from trans-eQTLs on the same chromosome arm. For this reason, in this paper we use an operational definition of trans-eQTLs as those acting across chromosomes, where they can be confidently distinguished. If desired, the 'set-to-zero' step can be skipped to include within-chromosome correlations in the snQTL analysis.

Next, we perform statistical tests to identify genetic markers affecting co-expression networks; see Step 1 in Fig 5.1. At each marker i, we test the null hypothesis:

$$H_0 \colon N_A = N_B = N_H. \tag{5.1}$$

In the next section, we will provide several test statistics based on the sparse multilinear spectral components of the correlation matrices. Let $D_{AB} = \hat{N}_B - \hat{N}_A$, $D_{AH} = \hat{N}_H - \hat{N}_A$, and $D_{BH} = \hat{N}_H - \hat{N}_B$ denote the three pairwise differential networks, where $\hat{\cdot}$ denotes the sample correlation matrix. The sparse multilinear spectral components of differential networks allow us to test for classical genetic dominance effects as well as a broad range of genetic effects onto the entire co-expression networks. We use permutation to obtain the p-value for the hypothesis test in Eq (5.1). The output is summarized as a Manhattan plot of association p-values across the genome. Our framework allows both traditional cases where the sample size is greater than the number of genes ($n > p$) and to high-dimensional cases where $n < p$. We address high-dimensionality by introducing sparsity into the test statistics and using permutation testing for a robust assessment of the null distribution.

Last, we estimate the joint differential network at the associated marker; see Step 2 in Fig 5.1. The joint differential network is a gene network with weighted edges that represent co-expressions affected by a genetic marker. Unlike the pair-

wise differential network, which describes the co-expression changes just between two genotypes (e.g., AA versus AB), the joint differential network summarizes prominent co-expression changes across three genotypes. We use sparse tensor decomposition to obtain the leading eigenvectors in the pairwise differential correlations. These eigenvectors summarize the differential signal into a single network. The resulting joint differential network has nodes representing genes, edges representing co-expression changes associated with the genetic marker, and edge weights and signs representing the magnitude and direction (enhancement or reduction) of these changes.

snQTL testing and joint differential network estimation via sparse tensor decomposition

We briefly introduce the sparse symmetric tensor decomposition (SSTD) in our contexts. Let $\mathcal{D} \in \mathbb{R}^{p\times p\times q}$ be an order-3 tensor with each of the q slides being a symmetric p-by-p matrix. We say $\mathcal{D}$ is sparse and of rank 1 if $\mathcal{D}$ satisfies the SSTD model:

$$\mathcal{D} = \Lambda v \circ v \circ u, \text{ or equivalently } \mathcal{D}_{jkl} = \Lambda v_j v_k u_l,$$

for all $(j,k,l) \in \{1,\ldots,p\}\times\{1,\ldots,p\}\times\{1,\ldots,q\}$, where $\circ$ denotes the vector outer product, v and u are norm-1 vectors in $\mathbb{R}^p$ and $\mathbb{R}^q$, respectively, and v is further sparse with $\|v\|_0 \leq R$ for some constant $R \leq p$, and $\Lambda \in \mathbb{R}_+$. Here $\|\cdot\|_0$ is the L_0 norm that counts the number of non-zero entries in the vector. The constraint on $\|v\|_0$ controls the sparsity on the first two modes, with a smaller R resulting in a sparser v. We call Λ, v, and u, the sparse leading tensor eigenvalue (sLTE), the sparse tensor eigenvector, and the loading vector, respectively.

In our snQTL framework, we define an order-3 differential tensor $\mathcal{D} \in \mathbb{R}^{p\times p\times 3}$ by stacking the three pairwise differential networks D_{AB}, D_{AH}, D_{BH} together. To summarize the signal in $\mathcal{D}$, we compute the SSTD approximation to the tensor $\mathcal{D}$. Specifically, we solve for the spectral components (Λ, v, u) that minimize the least

square approximation error

$$\min_{\substack{(\Lambda,v,u)\in\mathbb{R}_+\times\mathbb{R}^p\times\mathbb{R}^q,\\ \|v\|_2=\|u\|_2=1,\|v\|_0\leq R}} \|\mathcal{D}-\Lambda v\circ v\circ u\|_F^2, \tag{5.2}$$

where $\|\cdot\|_F$ denotes the Frobenius norm defined as the squared sum of tensor entries, and $\|\cdot\|_2$ denotes the vector L_2 norm. We denote the sLTE solution as $\Lambda(\mathcal{D})$, with $\mathcal{D}$ being the input differential tensor. The sLTE $\Lambda(\mathcal{D})$ represents the global strength of the co-expression changes across the three genotype groups. A larger sLTE suggests a stronger deviation from the null hypothesis in (5.1). Our test statistics, named $\text{Stat}_{\text{test}}$, is defined using the sLTE:

$$\text{Stat}_{\text{tensor}} = \Lambda(\mathcal{D}).$$

Our snQTL also features the estimation of a joint differential network. The sparse tensor eigenvector, $v = v(\mathcal{D})$, and loading vector, $u = u(\mathcal{D})$, together capture a lower-dimensional representation of $\mathcal{D}$. We call the leading matrix approximation, vv^T, the "joint differential network". This network captures the overall co-expression network changes in response to the genetic variation at the marker of interest. We call the element-wise squared eigenvector, denoted as v^2, the "gene leverage". The vector v reflects the contribution of genes to the differential network, with higher values indicating greater connectivity within the network. The loading vector, u, represents the weights for comparisons among the three genotype groups (e.g., AA-AB, AA-BB, AB-BB), with larger magnitudes indicating a greater contribution from each comparison.

Our snQTL is inspired from earlier work on SSTD Sun et al. (2017). However, we introduce key modifications that tailor SSTD to our specific needs in snQTL analysis. We explicitly considers the symmetry and sparsity in the first two modes of the tensor, making SSTD a better fit for our framework (details in *Materials and Methods*). Furthermore, unlike earlier work that focuses on the decomposition only Sun et al. (2017), our primary goal is hypothesis testing within the context of snQTL analysis. We have developed specific tools for this purpose.

snQTL testing via sparse matrix decomposition

We also propose an optional statistic for (5.1), based on extension of sparse leading matrix eigenvalue (sLME) Zhu et al. (2017). The sLME of a matrix D is defined as

$$\lambda(D) = \max_{v \in \mathbb{R}^p, \|v\|_2 = 1, \|v\|_0 \leq R} |v^T D v|. \tag{5.3}$$

The sLME represents the maximum eigenvalue of matrix D subject to the sparse eigenvectors. Our second test statistics, named "max", is defined as the maximal sLME from all three pairwise comparisons:

$$\text{Stat}_{\text{max}} = \max\{\lambda(D_{AB}),\ \lambda(D_{AH}),\ \lambda(D_{BH})\}.$$

Under the null hypothesis in (5.1), all pairwise differences (D_{AB}, D_{AH}, D_{BH}) are zero matrices, resulting in a zero max statistic. Conversely, a larger max statistic indicates higher differences in at least one pairwise comparison, making it well-suited for joint comparison of multiple networks.

Our max statistic generalizes the earlier work from pairwise comparison (Zhu et al., 2017) to joint comparison of multiple matrices. Other methods include L_2-type statistics Li and Chen (2012) that consider all entries in the comparison, and L_∞-type statistics Chang et al. (2017) that focus on the largest deviation. However, the L_2-type statistics assume all genes contribute equally, while the L_∞-type statistics capture only the single most extreme gene pair. In contrast, the spectral statistic, sLME, is well-suited for scenarios where the genetic effects are weak and sparse, meaning that a small subset of genes exhibit moderate effects (Zhu et al., 2017; Pantelis et al., 2014; Replication et al., 2014). This aligns with the biological expectation that genes might have significant but subtle co-expression changes. Additionally, the sparsity in sLME promotes result interpretability and faster computation.

Algorithm implementation

We design an iterative algorithm that alternatively updates the decomposition components to approximately solve (5.2). We adopt the penalized matrix decom-

position Zhu et al. (2017); Witten et al. (2009) to approximately solve for sLME in (5.3). In practice, we also consider variants of tensor and max statistics, such as the sum of sLMEs and the squared sLTE (*S1 Appendix*). More variants can be designed based on problem contexts. If the joint differential network is of interest, the tensor statistic is recommended, as the tensor approach facilitates network estimation. If the goal is solely snQTL testing, both tensor and matrix approaches perform similarly. For all test statistics, we use permutation to approximate the null distributions and obtain the empirical p-values. The number of permutations and the sparsity hyper-parameter R can be adjusted as needed. See *Materials and Methods* for more details. The code implementation of snQTL is provided at `https://github.com/Marchhu36/snQTL`.

Analysis of simulated data

In the simulation, we focus on the individuals drawn from an F2 hybrid generation derived from crosses between genetically divergent populations, with parent-of-origin diagnostic genetic markers spread across all chromosomes. This is to mimic the genotype patterns in the three-spined stickleback data. We also performed similar experiment using GWAS synthetic data; the simulation results are in the "S1 Appendix". In general, our approach is suitable for any other type of cross designs.

We first evaluated the efficiency of our snQTL framework on synthetic data for 200 genes across 20 chromosomes. We started with genetically divergent homozygous parents, and simulated the genotypes for an F1 cross and for an F2 intercross generation with random chromosomal crossing overs. For each F1 gamete, we simulated one recombination event per chromosome per gamete, randomly placed along the chromosome with a uniform distribution. The F2 hybrids' gene expression counts were generated from Poisson distributions with parameters varying by genotypes. We randomly selected one gene as the snQTL and generated three network effects associated with the snQTL. Then, we altered the expression levels of all 200 genes based on the additive network effect associated with this selected snQTL. In our current simulations, all 200 genes were considered candidate loci for

snQTL detection. In real scenarios, any marker linked to genetic variants – such as single nucleotide polymorphism (SNPs), indels, or haplotypes – could serve as candidate loci for snQTL detection. We tested the framework with varying hybrid population sizes from 50 to 500 to assess performance cover various scenarios. The detailed procedures for synthetic data generation are in *S1 Appendix*.

Fig 5.2 confirms the similarity between the synthetic and real F2 hybrid three-spined stickleback data Weber et al. (2022). The similar block diagonal patterns in the genetic correlation heatmaps (Fig 5.2A) suggest the LD among real and simulated markers. In both real and synthetic data, markers from the same chromosome show higher genotype correlations than those from different chromosomes. Our synthetic genotype generation captures this notable block structure. The overlapped histograms of expression counts (Fig 5.2B) validate our simulation procedures, indicating parameter values effectively mimicked real datasets.

We compared three methods on the synthetic data: the snQTL framework with max statistic, with tensor statistic, and a local approach based on F-tests for linear regressions of pariwise co-expression against genotypes. This local approach is similar to previous co-QTL analyses Van Der Wijst et al. (2018); Yuan et al. (2022a). We assessed both statistical power and type I error by applying all tests at the snQTL and non-snQTLs. Average test p-values and true positive (TP)/negative (TN) rates were recorded across 50 replicates for each population size.

Fig 5.2C demonstrates the superior statistical power of our snQTL framework, especially with larger populations. Our snQTL-based methods demonstrate a significant improvement as the sample size increases, while the local method shows only a minor improvement with additional samples. The out-performance suggests that the snQTL framework effectively addresses the multiple testing burden and tends to lead to more discoveries than the local approach. Additionally, the high TN rates at non-snQTLs support the high accuracy of the snQTL framework for snQTL identification. This outerperformance of snQTL is consistent across various simulation setups, including GWAS-like synthetic data, different sparsity parameters, and hybrid generations. For extra simulations, please refer to S1 Appendix.

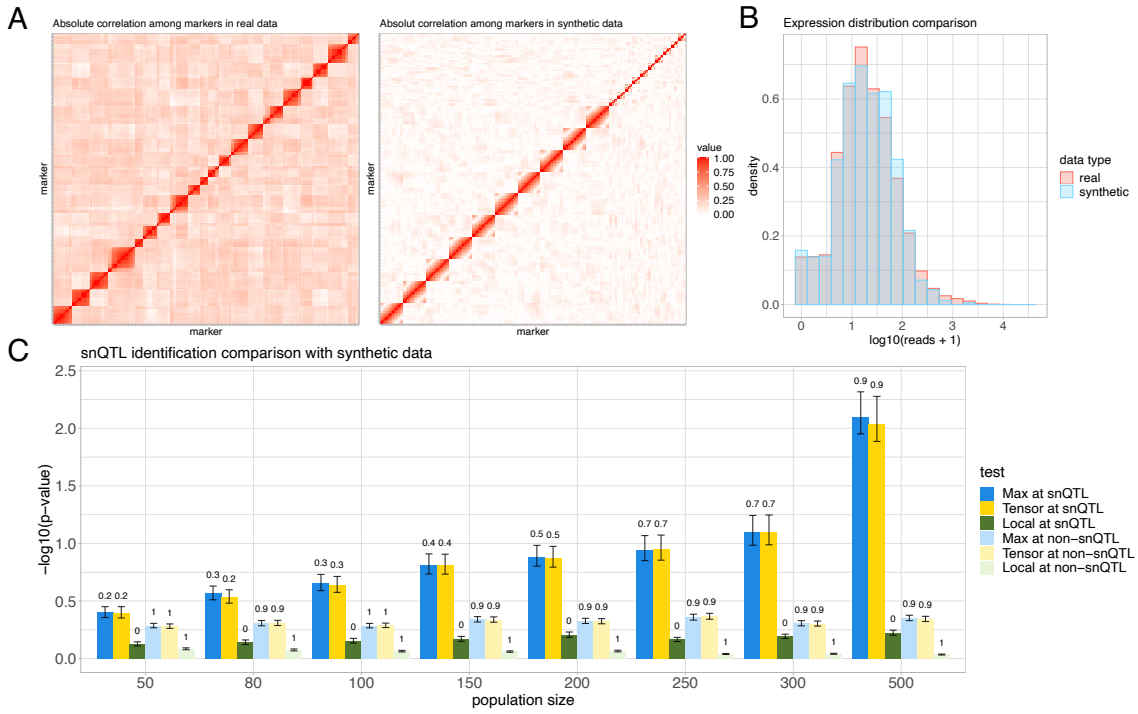

Figure 5.2: **Analysis of simulated data.** Synthetic datasets in three panels have the same parameter setup. (A) Absolute genetic correlation heatmaps among the markers in real F2 hybrid three-spined stickleback data Weber et al. (2022) and synthetic data. Markers are ordered following their positions on the genome. Genetic correlations are measured by absolute sample Pearson correlation coefficients between the genotypes of two markers. (B) Density histograms for expression counts in real stickleback and synthetic data. The parameters in synthetic data generation are carefully chosen to mimic the real data. (C) Barplots comparing the snQTL identification performances for snQTL framework and local method (F-test for regression of pairwise co-expression onto genotype) on synthetic data with varying population size from 50 to 500. We set sparsity parameter $R = p$ in snQTL for a fair comparison with the non-sparse local method. For results labeled "at snQTL", the y-axis is the observed $-log_{10}$(p-value) for tests at the single true snQTL; for results labeled "at non-snQTL", the y-axis is the averaged observed $-log_{10}$(p-value) from three tests at randomly selected non-snQTL markers. True positive (or negative) rates for the tests at snQTL (or non-snQTL) are shown above the bars. All reported numbers are averaged across 50 replications for each population size.

Performing snQTL to map stickleback loci affecting co-expression networks

We conducted snQTL analysis on the three-spined stickleback (*Gasterosteus aculeatus*) data Weber et al. (2022) to reveal the genetic landscape for co-expression networks in sticklebacks. These datasets are from a QTL mapping study in which wild fish were obtained from two lakes on Vancouver Island (Roberts Lake and Gosling Lake; RR and GG), and eggs/sperm mixed in petri dishes to generate F1 hybrids (RG). These hybrids were reared to maturity in an aquarium lab at the University of Texas and intercrossed to generate F2 intercross hybrids (RG*RG) and reciprocal backcrosses (RG*GG, GG*RG, RG*RR, RR*RG). Although hybrid crosses constituted a mixture of maternal backgrounds, maternal effects were excluded in our analyses. All F2 generation fish were reared to maturity in the laboratory and experimentally exposed to a cestode parasite, then euthanized 42 days post-exposure. Transcriptomic dataset was collected from head kidneys (pronephros, a major immune organ in fish) using Tag-Seq Lohman et al. (2017). The cross design, sequencing methods, and bioinformatics pipelines are described in depth in earlier work Weber et al. (2022); Fuess et al. (2021).

The raw dataset consists of gene transcript counts and genotypes for 234 markers, for 351 samples from F2 generations and backcrosses. The genetic markers in the stickleback data are biallelic SNPs obtained from ddRADseq (reduced representation genomic sequencing) of the parental stickleback, and the hybrids. The SNPs identified by ddRADseq were filtered to only include ancestry-informative markers that exhibit fixed differences between the two populations (e.g., one allele is only found in Roberts Lake, the alternate allele is only found in Gosling Lake). Details of the genotyping procedure and SNP calling for mapping are provided in Weber et al. (2022). We preprocessed the data with the following procedure. First, to eliminate non-functional variations, we normalized the read count matrix and regressed expressions against the sex and population covariates, retaining the residuals (*S1 Appendix*). Second, we focused the analysis on the top 10,000 genes with the highest adjusted mean expressions, as more information may be involved

with actively highly expressed genes. The number 10,000 was chosen to ensure computational efficiency. In general, our method allows more genes in the initial co-expression network analysis, since our method avoids multiple testing issues. Relaxing the filtering step could reveal important biological patterns, as it is unclear what expression levels lead to meaningful shifts in networks or phenotypes. Other filtering strategies, such as variance filtering and leveraging biological knowledge, may be helpful in other applications.

In addition, we considered the infection status of the sample fish as cestode infection is likely an environmental confounder. We added the worm presence as a predictor in the pre-processing regression step. Our snQTL analysis exhibited the same conclusions (*S1 Appendix*) before and after the additional procedure, suggesting the robustness of our discoveries to the infection status. For conciseness, we presented only the analysis without infection covariates in this paper. We leave further analyses involving more covariates and genes for future investigations.

Identification of stickleback snQTLs

We performed snQTL analysis on stickleback data using both tensor and max statistics. Both approaches lead to similar testing results (*S1 Appendix*), demonstrating the robustness of our snQTL identification. We present the findings using the tensor statistic here, as the tensor approach also facilitates joint differential network estimation. The Manhattan plot in Fig 5.3A shows 21 stickleback snQTLs concentrated at Chr 3, Chr 8, and Chr 18. This clustering pattern of snQTLs aligns with the LD structure among markers (Fig 5.2A). For instance, markers X77 and X75 exhibit highly correlated genotypes, leading both to be significant in snQTL tests. The three chromosomes of interest all exhibit extensive and stronger signals of snQTL associations compared to other chromosomes.

To further narrow down potential functional regions, we examined within each snQTL region for coding genes with strong genomic signatures of past natural selection. Specifically, we used published population genomic data: allele frequency estimates obtained from PoolSeq of $\sim$ 100 fish from each of three populations

(Roberts Lake, Gosling Lake, and a marine outgroup). We calculated population branch statistics (PBS) measuring accelerated evolution in each lake (Roberts or Gosling), relative to an ancestral marine population (Sayward), as described in earlier work Weber et al. (2022). Large PBS in either lake population indicates a gene that was likely a target of natural selection within the lake in question, since its colonization $\sim$ 12,000 years ago.

Several protein-coding genes lie in regions adjacent to PBS outliers within snQTLs (*S1 Appendix*). We focused our analysis on genes near the largest snQTL on Chr 18 (Fig 5.3). None of these genes harbored coding variants but two were represented in our expression data: cellular communication network factor 6 (*ccn6*) and laminin subunity alpha 4 (*Lama4*). Although the *Lama4* expression differs little between parental populations, the *ccn6* expression was significantly lower in Gosling fish ($t = 2.115, df = 97.886, p = 0.037$, Fig 5.3D). In addition, *ccn6* is differentially expressed in our genetic crosses ($p = 0.0511$, *S1 Appendix*), but we found no evidence of protein coding changes between any of the populations. The gene *ccn6*, also known as *wisp3*, has 4 distinct protein domains that perform distinct functions Perbal (2018), several of which have notable connections to the stickleback system. Secreted *ccn6* can bind to and limit insulin growth factor-1 (*igf-1*) signaling, thereby suppressing cell growth and metabolic potential Repudi et al. (2013), as well as mediating fibrotic responses Yeger and Perbal (2016); Song et al. (2023). The gene *ccn6* also acts as a transcription factor that activates genes necessary for formation of the mitochondrial electron transport system Padhan et al. (2020) and indirectly regulates reactive oxygen species (ROS) levels Miller and Sen (2007). Gosling fish produce significantly less ROS, display less cestode-induced fibrosis, and grow faster than Roberts fish. It is worthy noting that in humans, the *ccn6* expression is largely restricted to kidney, skin and testes, consistent with an organ-specific regulatory role Weber et al. (2022); Fagerberg et al. (2014).

Joint differential network at snQTL locus X419

We further estimated joint differential networks for the significant snQTLs identified in our snQTL analysis. We found that most snQTLs are associated with similar sets of genes with high leverages, resulting in joint differential networks with comparable patterns (*S1 Appendix*). Here, we present the joint differential network at the most significant snQTL, X419 on Chr 18. We ranked genes based on their leverage scores from our method. We found that the top 10 genes achieved a cumulative leverage of 0.54, and the top 100 genes achieved a cumulative leverage of 0.9. We called the top 10 genes with highest leverages the “primary genes", and the remaining top genes the “secondary genes". These top 100 genes distribute widely on the genome, from the scaffold region and mitochondrial genome (MT) to all chromosomes (Fig 5.4A). This wide distribution of top genes implies the capacity of snQTLs to impact co-expressions throughout the whole genome. Such cross-chromosome influences are likely to represent functional genotype-network associations. None of the primary genes were identified in the previous differential expression (DE) analysis of pure parental populations (Lohman et al., 2017), which highlighting their unique roles in the co-expression network effects. In addition, the loading values for the genotype comparisons GG-RG and RG-RR are 0.498 and 0.31, respectively. The result suggests that the co-expressions between primary and secondary genes, except those with *bbx* and *otog*, are reduced in Gosling Lake fish and enhanced in Roberts Lake fish (Fig 5.4B). Moreover, the loading and co-expression networks for three genotypes (Fig 5.4C-E) show that differential networks for GG-RG and for RG-RR are comparable, indicating the nearly additive genetic effects to the co-expression networks.

We found that most genetic co-expression variations occur between a primary gene and a secondary gene and between two primary genes. (Fig 5.4B). We note that many of the primary genes (*hbae5*, two *hbe1* paralogs, and the novel gene *ENSGACT00000018413*, which is orthologous to *hba2* in other species of fish) are hemoglobin subunits expressed in red blood cells and directly participate in oxygen transport activities, while the others are involved in closely related biological pro-

cesses, such as blood vessel development (*hsp90ab1*) and carbohydrate metabolism (*otog*) (Table 5.1). These functions are consistent with decreased expression of *ccn6* being connected to elevated rates of *igf-1* signaling and cell replication in the head kidney, which is the hematopoietic organ in fish. Similarly, overexpression of heat shock proteins (i.e., *hsp90ab1*) can be stimulated either via pharmacological suppression of *igf-1* Wan et al. (2015) or dysregulation of the electron transport chain in mitochondria, which is another major function of *ccn6*. Although the precise role of *mmp16b* has not been well characterized, *igf-1* is connected to the expression of other *mmp*s. Our analysis demonstrates the power of snQTL framework with functional annotation for unraveling the genetic basis of co-expression networks.

Table 5.1: **List of primary genes with top 10 leverage scores in joint differential network at X419 on Chr 18.**

Transcript ID	Leverage	Gene	Chr	Gene and Protein GO annotations
ENSGACT00000019169	0.1433	hbae5	Scaffold 112	heme, iron ion, oxygen binding; oxygen carrier activity
ENSGACT00000018425	0.1271	hbe1	11	heme, iron ion, oxygen binding; oxygen carrier activity
ENSGACT00000026622	0.0765	bbx	7	DNA binding
ENSGACT00000018413	0.0656	-	11	heme, iron ion, oxygen binding; oxygen carrier activity
ENSGACT00000026589	0.0263	-	7	uncharacterized
ENSGACT00000022959	0.0232	otog	2	carbohydrate metabolic process
ENSGACT00000027730	0.0222	cox2	MT	copper ion binding; cytochrome-c oxidase activity in mitochondrion & respirasome
ENSGACT00000017921	0.0214	hsp90ab1	18	blood vessel development; leukocyte migration; response to estrogen
ENSGACT00000017116	0.0187	-	8	serine-type endopeptidase activity; proteolysis
ENSGACT00000018389	0.0185	hbe1	11	heme, iron ion, oxygen binding; oxygen carrier activity

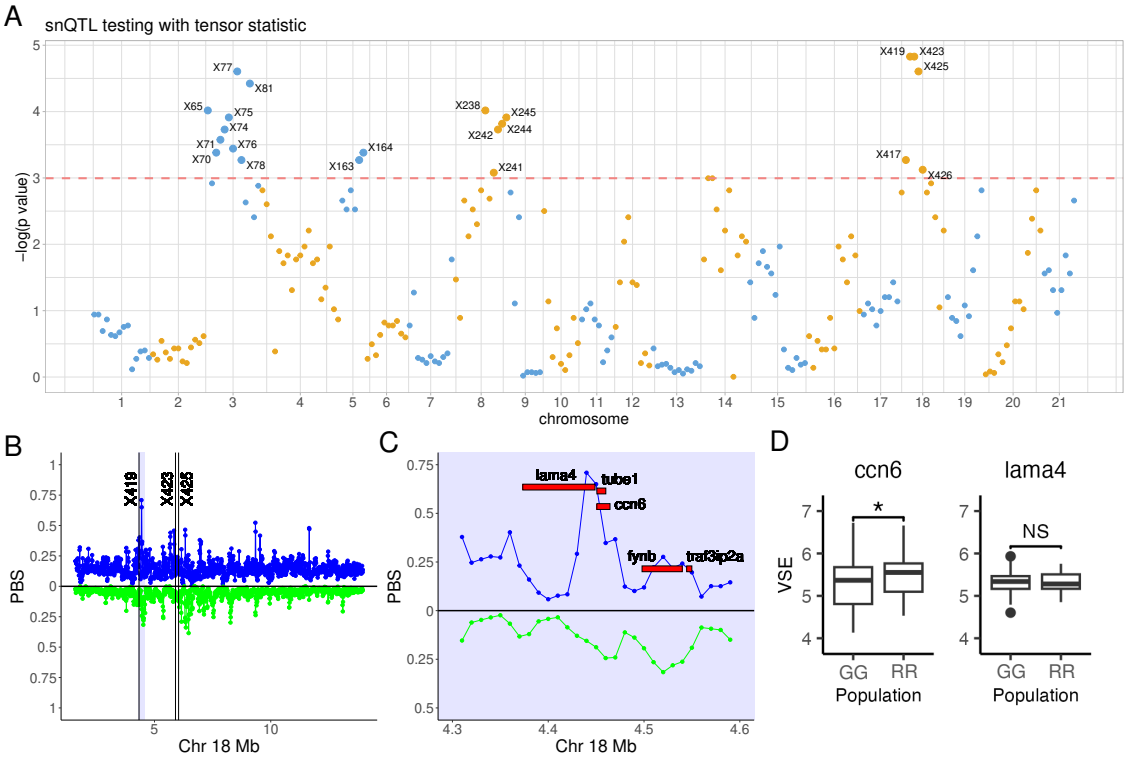

Figure 5.3: **Identification of stickleback snQTLs via snQTL framework.** (A) Manhattan plot for snQTL testing with tensor statistics marks 21 stickleback snQTLs, mainly clustered in Chr 3, Chr 8, and Chr 18. The y-axis represents the natural logarithms of p-values. The snQTLs are deemed with testing p-values smaller than 0.05 (above the dashed line). (B) Strong genomic targets of selection with high population branch statistic (PBS) distribute around the outstanding snQTLs (markers X419, X423, and X425) in Chr 18. Values above the medial line represent higher PBS in Gosling Lake (blue); values below the line represent higher PBS in Roberts Lake (green). (C) Zoomed-in shadowed area in (B). Development regulation genes, lama4 an d ccn6, locate tightly around marker X419 with high selection speed. (D) Variance stabilized expressions (VSE) for ccn6 and lama4 in Gosling (GG) and Roberts (RR) lakes.

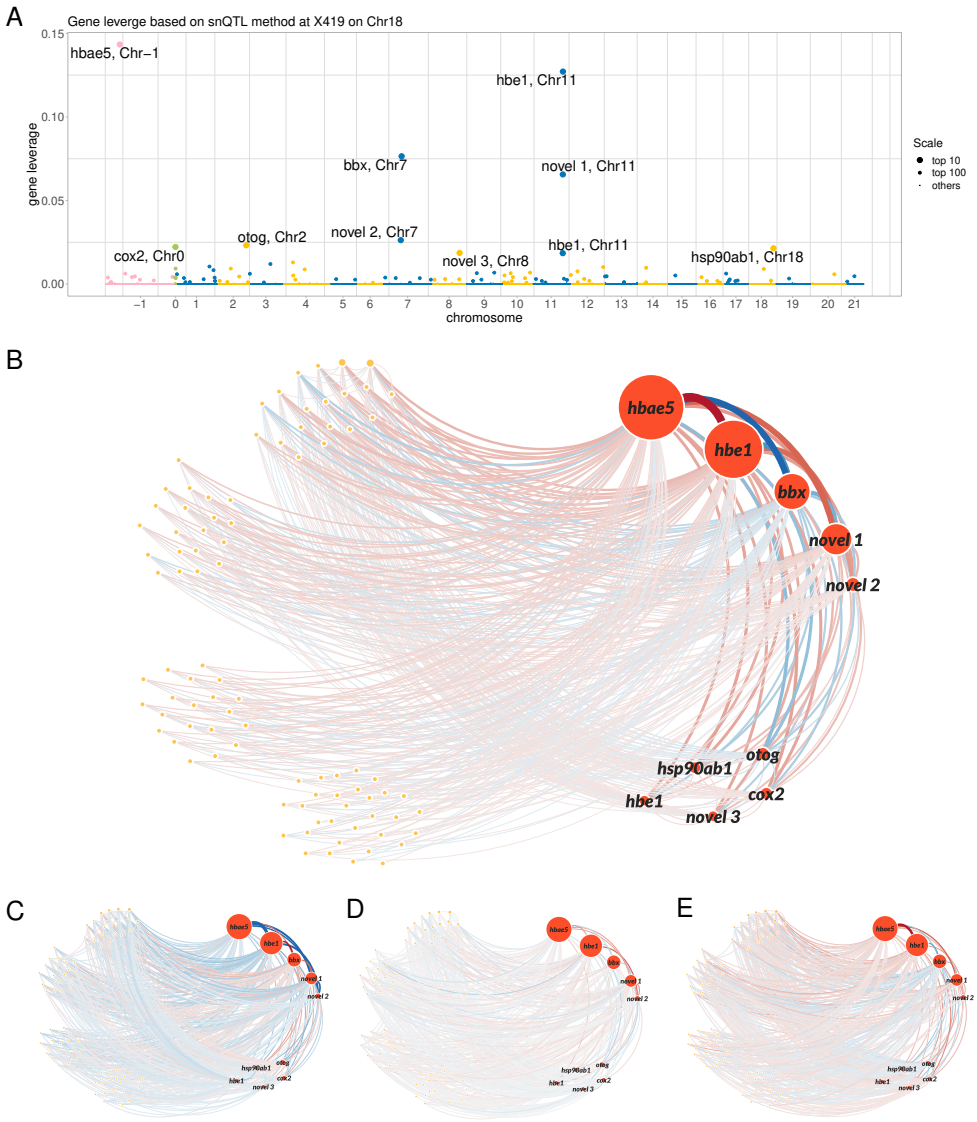

Figure 5.4: **Joint differential network analysis at snQTL X419 on Chr 18.** (A) Leverage scores for 10000 genes. Primary genes with top 10 leverage are highlighted with transcription IDs. Mitochondrial genome (MT) and scaffold region are coded as Chr 0 and Chr -1, respectively. (B-E) Networks for primary (red annotated nodes) and secondary (orange nodes) genes with top 100 leverages. The edge width indicates the connection strength between two genes; the diameter of node indicates the leverage of the gene; the color indicates enhancement (red) or reduction (blue) of the connection compared with average level. (B) Joint differential network at X419 with top 10% strongly connected edges. A wider edge implies a stronger genetic variation in the co-expression of the gene pair. Most genetic co-expression variations occur between the primary and secondary genes. (C-E) co-expression networks corresponding to the genotypes GG, RG, and RR at X419, respectively. The linear changes in the colors of edges imply the nearly additive genetic effect to the co-expression networks. novel 1: ENSGACT00000018413; novel 2: ENSGACT00000026589; novel 3: ENSGACT00000017116.

5.3 Discussion

Gene co-expression networks play a pivotal role in translating genotype into phenotype. This suggests that phenotypic evolution may often be a consequence of evolution not just of single genes' protein structure or expression level, but also by changes of co-expression patterns among genes Oldham et al. (2006); Ovens et al. (2021); Crow et al. (2022). For gene co-expression networks to evolve, there must be genetic variations within species that impact the network structure, which selection (or drift) might act on. Therefore, there is a need for methods capable of identifying loci (or chromosomal regions) that are associated with changes in co-expression networks. While methods exist for analysing pairwise gene co-expression Van Der Wijst et al. (2018); Baker et al. (2019); Yuan et al. (2022a); Li et al. (2023); Wei et al. (2022); Kolberg et al. (2020), a key challenge lies in methods that can analyze gene co-expression across entire networks.

Methodological significance

Our snQTL framework offers a methodological advance in network-based association study. Unlike traditional co-QTL methods that test millions of gene pairs independently, snQTL treats the entire co-expression network as a single entity. This dramatically reduces the multiple testing burden. Furthermore, snQTL leverages a tensor spectral statistic that captures the overall signal across the entire network. This approach avoids the need for pre-selecting candidate gene pairs, which can introduce bias. Additionally, unlike regression-based methods that assume an additive genetic effect, snQTL allows for a broad range of genetic effects. The flexibility enables the detection of snQTLs as long as a significant difference exists in co-expression network between genotypes.

The power of snQTL extends beyond co-expression networks. The framework can be generalized to analyze various networks, including microbial networks, proteomic networks, and others. With minor adjustments, snQTL can also handle directed networks like transcription factor binding networks and metabolic network. The core idea of snQTL can be applied for general mapping tasks beyond

genetics. For example, the method can handle comparisons of more than three networks, allowing investigation of associations with various discrete factors, such as treatment, location, or environmental conditions.

Several future improvements can be made to snQTL. Currently, snQTL removes all within-chromosome co-expression, since between-chromosome co-expressions are less likely to be generated by LD effects and more likely to indicate functional connections. Future improvements could incorporate recombination maps to identify unlinked markers on the same chromosome and linked markers on different chromosomes, providing a more biologically relevant approach. The other potential extension is on the use of SSTD. The current rank-1 SSTD approximation in snQTL captures the strongest signal in the network difference. Extending this to a higher-rank model could reveal more delicate signals, potentially leading to additional discoveries.

Biological significance

One of the “grand challenges" of biology is to understand the details of how genotypes produce phenotypes, and thereby develop tools to predict phenotypes. Genotype-phenotype prediction remains a challenge because most phenotypes are the emergent result of complex interactions between numerous genes. Network analyses offer a promising toolkit for representing these complex interactions. Such tools have been applied to gene-gene co-expression data Zhang and Horvath (2005); Zhao et al. (2010); Farhadian et al. (2021), single-cell RNAseq data Wang et al. (2021), gene-gene epistasis effects Costanzo et al. (2016), proteomic data Chisanga et al. (2017), and beyond, with the goal of describing the logic of genetic regulatory “circuits". The hope is that this network-based approach can reveal rules of life not visible for single genes and their mRNA and protein products, or simple pairwise gene interactions.

Our snQTL analysis of three-spined stickleback gene expression illustrates this potential benefit. We identified three chromosomes with significant snQTLs. Using population genomic data, we were able to identify a candidate gene under especially

strong selection within the snQTLs on Chr 18. The gene *ccn6* is a highly pleiotropic gene known to affect growth, metabolism, fibrosis, ROS production, and hence with great potential for network-wide effects in the immune organ sampled for transcriptomics. It appears likely that *ccn6*-mediated changes in electron transport chain function is affecting ROS production differences previously documented between the hybridized populations, with additional consequences for a protective fibrosis phenotype. This gene was not flagged in prior differential expression analyses of the same dataset. Although *ccn6* is expressed at significantly lower levels in Gosling than Roberts Lake fish, the differential expression was not exceptionally large. In contrast, the snQTL (aided by selection scans) makes this gene an important candidate for multivariate phenotypic effects. This result highlights a major limitation in how we currently search for expression-related evolutionary differences: we are most likely to focus on individual loci with large shifts in expression. However, even small changes in expression of one gene can be amplified via downstream effects of entire networks of genes, thereby exerting large phenotypic effects. Scanning large expression networks for correlated changes holds a great promise for uncovering evolving genes whose expression is either highly noisy with respect to genotype, or whose expression is only moderately shifted across populations.

Taken together, our snQTL analysis offers a powerful, effective, and adaptable framework for mapping QTLs that affecting network-based co-expression. We believe our approach brings a broad impact to the genetics community.

5.4 Materials and Methods

Sparse matrix decomposition

We use penalized matrix decomposition Zhu et al. (2017) to approximately solve for sparse symmetric matrix decomposition in (5.3). The PDM with input matrix D is expressed as

$$\max_{\|v\|_2 \leq 1, \|v\|_1 \leq \sqrt{R}} \mathrm{tr}(D(vv^T)). \tag{5.4}$$

By Zhu et al. (2017), the solutions to (5.4) always have $\|v\|_2 = 1$ and satisfy the inequality $\|v\|_1^2 \leq \|v\|_0 \leq R$. Therefore, (5.4) is an good approximation to sLME in (5.3). We follow the algorithm in Zhu et al. (2017) to solve (5.4).

Sparse symmetric tensor decomposition algorithm

We solve the optimization problem (5.2) via SSTD by an iterative algorithm. For a tensor $\mathcal{D} \in \mathbb{R}^{p_1 \times p_2 \times p_3}$ and vectors $v^{(k)} \in \mathbb{R}^{p_k}$ for $k = 1, 2, 3$, we define the tensor-by-vector product on mode 1, mode 2, and mode 3 as

$$\begin{aligned}
\mathcal{D} \times_1 v^{(1)} &= \sum_{i=1}^{p_1} v_i^{(1)} \mathcal{D}_{i::}, \\
\mathcal{D} \times_2 v^{(2)} &= \sum_{i=1}^{p_2} v_i^{(2)} \mathcal{D}_{:i:}, \\
\mathcal{D} \times_3 v^{(3)} &= \sum_{i=1}^{p_3} v_i^{(3)} \mathcal{D}_{::i}.
\end{aligned}$$

Given input tensor $\mathcal{D}$, our decomposition algorithm is presented as follows:

1. *Input.* Differential tensor $\mathcal{D} \in \mathbb{R}^{p \times p \times 3}$, sparsity parameter R, and iteration number T.

2. *Initialization.* Randomly initilize the unit vectors $v^{(0)} \in \mathbb{R}^p, u^{(0)} \in \mathbb{R}^3$.

3. *For iteration* $t = 1, \ldots, T$, alternatively update the decomposition components $v^{(t)}$ and $u^{(t)}$:

$$\begin{aligned}
v^{(t)} = &\underset{\|v\|_2 \leq 1, \|v\|_1 \leq \sqrt{R}}{\operatorname{argmin}} \operatorname{tr}(D^{(t)}(vv^T)), \\
\text{with } D^{(t)} &= \mathcal{D} \times_3 u^{(t-1)}
\end{aligned} \tag{5.5}$$

   and

$$u^{(t)} = \text{Normalize}(\mathcal{D} \times_1 v^{(t)} \times_2 v^{(t)}).$$

4. *Output.* Output the eigen components $v(\mathcal{D}) = v^{(T)}, u(\mathcal{D}) = u^{(T)}$ and estimated sLTE

$$\Lambda(\mathcal{D}) = \mathcal{D} \times_1 v(\mathcal{D}) \times_2 v(\mathcal{D}) \times_3 u(\mathcal{D}).$$

Here Normalize$(v) = v/\|v\|_2$ denotes the vector normalization step. We make two comments on our algorithm. Previous work Sun et al. (2017) enforces by value truncation. In contrast, our approach achieves sparsity through an optimization process called PMD (Proximal Minimization with Duality) during the update of a variable $v^{(t)}$ in (5.5). Our approach is computationally faster and reflects the symmetry in our SSTD model. Second, in our construction of differential tensor input $\mathcal{D}$, the third slide D_{BH} can be expressed as the sum of first two slides D_{AB} and D_{AH}. While this linear relationship does not affect the final results of association testing, we choose to analyze the model using a full 3-layer tensor $\mathcal{D}$ for easier interpretation.

Permutation and empirical p-values

We used permutation to obtain empirical p-values based on our proposed test statistics. Specifically, at each marker, we repetitively shuffle three genotypes of samples, re-divide the expression dataset into three groups, and re-calculate the test statistics for B times. Let S denote the test statistic with original genotype, and S_b denote the test statistic with shuffled dataset in the b-th permutation for $b = 1, \ldots, B$. We obtain the empirical p-value as

$$\text{p-value} = \frac{1}{B} \sum_{b=1}^{B} \mathbb{I}\{S_b \geq S\},$$

where $\mathbb{I}\{\cdot\}$ is the indicator function.

In our stickleback data analysis, we first obtained the empirical p-values for all markers with $B_0 = 100$ permutations for preliminary snQTL screening. For the markers showing preliminary empirical p-values smaller than 0.05, we re-ran the tests with $B = 500$ permutations for accurate p-values estimations.

Sparsity hyperparameter

We discuss the tuning of the sparsity hyperparameter R. A higher R leads to denser connection in the estimated joint differential network. Common methods for selecting R include cross-validation via random test-train splits. Other works Zhu et al. (2017); Witten et al. (2009) suggested setting R proportional to the feature dimension p. We recommend exploring a range of R and choose the one that balances the number of discoveries with their biological relevance.

For simulations in Fig 5.2C, we use a non-sparse setting ($R = p$) for a fair comparison with the non-sparse local method. Additional experiments with varying sparsity parameter are detailed in *S1 Appendix*. Our snQTL method shows a stable performance across a wide range of R. For the stickleback data analysis, we set the sparsity parameter R to $0.09p$. This aligns with the expectation that only a few thousand genes contributing to the main co-expression differences. In other applications, we recommend choosing R based on background knowledge and the scientific context.

In addition, we propose a data-driven approach for selecting R when no prior information is available. While our snQTL objective function (5.2) is not built on probabilistic distributional assumptions, following earlier works in similar contexts Sun et al. (2017); Allen et al. (2014), we suggest using a Bayesian Information Criterion (BIC)-based approach for parameter tuning. Given the differential tensor $\mathcal{D} \in \mathbb{R}^{p\times p\times 3}$, we choose the sparsity parameter R that minimizes the following BIC:

$$R_{\text{select}} = \underset{1\le R\le p}{\arg\min}\, 3p^2 \log\left(\frac{\|\mathcal{D} - \Lambda^R v^R \circ v^R \circ u^R\|_F^2}{3p^2}\right) + (2 + 2\|v^R\|_0)\log 3p^2, \quad (5.6)$$

where (Λ^R, v^R, u^R) are the SSTD components with the sparsity parameter R. The first term in (5.6) represents the squared loss of the SSTD decomposition, and the second terms is the penalty proportional to the number of parameters.

This BIC metric (5.6) balances the approximation error and sparsity, and it has been widely used for hyperparameter selection in tensor literature Sun et al. (2017); Allen (2012); Wang and Zeng (2019); Han et al. (2022a); Hu and Wang (2023).

We performed simulations to show the efficacy of the proposed approach; see *S1 Appendix* for details. We also applied the BIC selection to the real stickleback data. While the selected sparsity parameters R for some markers exceed $0.09p$, all testing results and joint differential network estimation conclusions remain the same across different choices of R. This consistency indirectly verifies the robustness of our current stickleback analyses. See *S1 Appendix* for additional simulation results and stickleback data analysis.

6 CONCLUSION AND FUTURE WORKS

In this thesis, we propose four statistically rigorous and computationally efficient methods for real tensor data analysis, each contributing to methodological development from distinct perspectives.

First, the SupTD method addresses the critical challenge of modeling complex interdependencies within tensor data and across multiple features. Applications to brain connectivity and political networks yield interpretable results, underscoring the practical value of SupTD. Moreover, SupTD lays a foundation for several extensions, including statistical tensor inference, mixed-effect tensor regression, and advanced decomposition models with non-trivial covariance structures.

Second, we comprehensively investigate multiway clustering under the degree-corrected tensor block model (dTBM). dTBM allows the node degree heterogeneity within clusters, which provides great flexibility in clustering models, but the extra complexity poses significant challenges in both statistics and computation. We establish the theoretical properties of dTBM and develop an efficient algorithm with practical utility and broad applicability.

Third, the LS-TBM framework addresses computational limitations of existing tensor clustering methods by decomposing the high-dimensional TBM algorithm into two low-dimensional, efficient steps. Theoretical analyses demonstrate that LS-TBM significantly reduces computational complexity while maintaining strong clustering performance. Empirical results confirm its substantial practical advantages and potential for broader applications in large-scale tensor analysis.

Finally, the snQTL framework advances network-based association studies in genetics by identifying genetic loci that influence global co-expression networks—overcoming key limitations of traditional QTL approaches. Applications to stickleback data highlight the biological significance of snQTL, emphasizing its potential to uncover novel gene-trait relationships. We envision snQTL as a versatile tool for network-based genetic analysis, with promising applications beyond the current study.

A APPENDIX OF CHAPTER 2

A.1 Proofs

Proof of Theorem 2.2

We denote several quantities:

$$\underline{\gamma} = \prod_{k\in[K]} \sigma_{\min}(\boldsymbol{X}_k), \quad \bar{\gamma} = \prod_{k\in[K]} \sigma_{\max}(\boldsymbol{X}_k), \quad \lambda = \min_{k\in[K]} \sigma_{\min}(\text{Unfold}_k(\mathcal{B}_{\text{true}})), \quad \text{(A.1)}$$

where $\underline{\gamma}$ quantifies the rank non-deficiency of feature matrices, $\bar{\gamma}$ quantifies the magnitude of feature matrices, and λ is the smallest singular value of mode-k unfolded matrices $\text{Unfold}_k(\mathcal{B}_{\text{true}})$ for all possible $k \in [K]$. For notational convenience, we drop the subscript $\mathcal{Y}$ from the objective $\mathcal{L}_{\mathcal{Y}}(\cdot)$ and simply write as $\mathcal{L}(\cdot)$. We write $\mathcal{L}(\mathcal{B})$ in place of $\mathcal{L}(\mathcal{C}, \boldsymbol{M}_1, \ldots, \boldsymbol{M}_K)$ when we want to emphasize the role of $\mathcal{B}$.

Proposition A.1 (sub-Gaussian residuals)**.** *Define the residual tensor $\mathcal{E} = [\![\varepsilon_{i_1,\ldots,i_K}]\!] = \mathcal{Y} - b'(\Theta) \in \mathbb{R}^{d_1\times\cdots\times d_K}$. Under the Assumption A2, $\varepsilon_{i_1,\ldots,i_K}$ is a sub-Gaussian random variable with sub-Gaussian parameter bounded by ϕU, for all $(i_1, \ldots, i_K) \in [d_1]\times\cdots\times[d_K]$.*

Proposition A.2 (Properties of tensor GLM)**.** *Consider tensor GLMs under Assumption A2.*

(*a*) (*Strong convexity*) *For all $\mathcal{B}$ and all realized data tensor $\mathcal{Y}$,*

$$\mathcal{L}(\mathcal{B}_{\text{true}}) \geq \mathcal{L}(\mathcal{B}) + \langle \nabla\mathcal{L}(\mathcal{B}_{true}), \mathcal{B}_{\text{true}} - \mathcal{B}\rangle + \frac{1}{2}\underline{\gamma}^2 L\|\mathcal{B}_{\text{true}} - \mathcal{B}\|_F^2,$$

where $\nabla L(\cdot)$ denotes the derivative of $\mathcal{L}$ with respect to $\mathcal{B}$.

(*b*) (*Model complexity*) *Suppose $\mathcal{Y}$ follows generalized tensor model with parameter $\mathcal{B}_{\text{true}}$. Then, with probability at least $1 - \exp(-p)$,*

$$Err_{ideal}(\boldsymbol{r}) := \sup_{\|\mathcal{B}\|_F=1, \mathcal{B}\in\mathcal{P}(\boldsymbol{r})} \langle \nabla\mathcal{L}(\mathcal{B}_{true}), \mathcal{B}\rangle \lesssim \bar{\gamma}\sqrt{\phi U(r^K + Kpr)}. \quad \text{(A.2)}$$

The proofs of Propositions A.1-A.2 are in Section A.1.

Proof of Theorem 2.2. First we prove the error bound for $\hat{\mathcal{B}}_{\text{MLE}}$. By the definition of $\hat{\mathcal{B}}_{\text{MLE}}$, $\mathcal{L}_{\mathcal{Y}}(\mathcal{B}_{\text{true}}) - \mathcal{L}_{\mathcal{Y}}(\hat{\mathcal{B}}_{\text{MLE}}) \leq 0$. By the strong convexity in Proposition A.2,

$$0 \geq \mathcal{L}_{\mathcal{Y}}(\mathcal{B}_{\text{true}}) - \mathcal{L}_{\mathcal{Y}}(\hat{\mathcal{B}}_{\text{MLE}}) \geq \langle \nabla \mathcal{L}(\mathcal{B}_{\text{true}}), \mathcal{B}_{\text{true}} - \hat{\mathcal{B}}_{\text{MLE}} \rangle + \frac{1}{2}\underline{\gamma}^2 L \|\mathcal{B}_{\text{true}} - \hat{\mathcal{B}}_{\text{MLE}}\|_F^2. \tag{A.3}$$

Rearranging (A.3) gives

$$\|\hat{\mathcal{B}}_{\text{MLE}} - \mathcal{B}_{\text{true}}\|_F \leq \frac{2}{\underline{\gamma}^2 L} \left\langle \nabla \mathcal{L}(\mathcal{B}_{\text{true}}), \frac{\hat{\mathcal{B}}_{\text{MLE}} - \mathcal{B}_{\text{true}}}{\|\hat{\mathcal{B}}_{\text{MLE}} - \mathcal{B}_{\text{true}}\|_F} \right\rangle \leq \frac{2}{\underline{\gamma}^2 L} \text{Err}_{\text{ideal}}(2\boldsymbol{r}),$$

where the last inequality comes from the definition of $\text{Err}_{\text{ideal}}(2\boldsymbol{r})$ and the fact that $\text{rank}(\hat{\mathcal{B}}_{\text{MLE}} - \mathcal{B}_{\text{true}}) \leq \text{rank}(\hat{\mathcal{B}}_{\text{MLE}}) + \text{rank}(\mathcal{B}_{\text{true}}) \leq 2\boldsymbol{r}$. By (A.2) in Proposition A.2, we have

$$\|\hat{\mathcal{B}}_{\text{MLE}} - \mathcal{B}_{\text{true}}\|_F \lesssim \frac{\bar{\gamma}\sqrt{\phi U}}{\underline{\gamma}^2 L} \sqrt{r^K + Kpr}, \tag{A.4}$$

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

Now, we specialize $\bar{\gamma}/\underline{\gamma}^2$ in the following two cases of assumptions on feature matrices.

[Case 1] Under Assumption A1 with scaled feature matrices, we have

$$\frac{\bar{\gamma}}{\underline{\gamma}^2} \leq \frac{c_2^K d^{K/2}}{c_1^{2K} d^K} \lesssim \sqrt{\frac{1}{d^K}}. \tag{A.5}$$

[Case 2] Under Assumption A1' with original feature matrices, the asymptotic behavior of extreme singular values (Rudelson and Vershynin, 2010) are

$$\sigma_{\min}(\boldsymbol{X}_k) \asymp \sqrt{d} - \sqrt{p} \text{ and } \sigma_{\max}(\boldsymbol{X}_k) \asymp \sqrt{d} + \sqrt{p}, \quad \text{for all } k \in [K].$$

In this case, we obtain

$$\frac{\bar{\gamma}}{\underline{\gamma}^2} \asymp \frac{(\sqrt{d} + \sqrt{p})^K}{(\sqrt{d} - \sqrt{p})^{2K}} \lesssim \sqrt{\frac{1}{d^K}}. \tag{A.6}$$

Combining (A.4) with either (A.5) or (A.6), in both cases we obtain the same conclusion

$$\|\hat{\mathcal{B}}_{\text{MLE}} - \mathcal{B}_{\text{true}}\|_F^2 \lesssim \frac{\phi(r^K + Kpr)}{d^K}. \tag{A.7}$$

Now we prove the bound for sinΘ distance. We unfold tensors $\mathcal{B}_{\text{true}}$ and $\hat{\mathcal{B}}_{\text{MLE}}$ along the mode k and obtain $\text{Unfold}_k(\mathcal{B}_{\text{true}})$ and $\text{Unfold}_k(\hat{\mathcal{B}}_{\text{MLE}})$. Notice that $\boldsymbol{M}_{k,\text{true}}$ and $\hat{\boldsymbol{M}}_{k,\text{MLE}}$ span the top-r left singular spaces of $\text{Unfold}_k(\mathcal{B}_{\text{true}})$ and $\text{Unfold}_k(\hat{\mathcal{B}}_{\text{MLE}})$, respectively. Applying Proposition A.3 to this setting gives

$$\sin\Theta(\boldsymbol{M}_{k,\text{true}}, \hat{\boldsymbol{M}}_{k,\text{MLE}}) \leq \frac{\|\text{Unfold}_k(\hat{\mathcal{B}}_{\text{MLE}}) - \text{Unfold}_k(\mathcal{B}_{\text{true}})\|_F}{\sigma_{\min}(\text{Unfold}_k(\mathcal{B}_{\text{true}}))} = \frac{\|\hat{\mathcal{B}}_{\text{MLE}} - \mathcal{B}_{\text{true}}\|_F}{\lambda}. \tag{A.8}$$

The proof is complete by combining (A.7) and (A.8). □

Proofs of Proposition 2.4 and Theorem 2.3

Proof of Proposition 2.4. We express the Gaussian model as

$$\mathcal{Y} = \mathcal{B}_{\text{true}} \times \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K\} + \mathcal{E},$$

where $\mathcal{E}$ is a noise tensor consisting of i.i.d. entries from $N(0, \sqrt{\phi})$. By QR decomposition on feature matrices, $\boldsymbol{X}_k = \boldsymbol{Q}_k \boldsymbol{R}_k$ for all $k \in [K]$, we have

$$\bar{\mathcal{Y}} = \mathcal{B}_{\text{true}} \times \{\boldsymbol{R}_1, \ldots \boldsymbol{R}_K\} + \bar{\mathcal{E}}, \tag{A.9}$$

where $\bar{\mathcal{Y}} = \mathcal{Y} \times \{\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_K\}$ and $\bar{\mathcal{E}} = \mathcal{E} \times \{\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_K\}$. Notice that entries of $\bar{\mathcal{E}} \in \mathbb{R}^{p \times \cdots \times p}$ are i.i.d drawn from $N(0, \sqrt{\phi})$ by the orthonormality of $\{\boldsymbol{Q}_k\}_{k=1}^K$. Reparameterize the signal in (A.9) as

$$\begin{aligned} \mathcal{S}_{\text{true}} := \mathcal{B}_{\text{true}} \times \{\boldsymbol{R}_1, \ldots \boldsymbol{R}_K\} &= \mathcal{C}_{\text{true}} \times \{\boldsymbol{R}_1 \boldsymbol{M}_{1,\text{true}}, \ldots \boldsymbol{R}_K \boldsymbol{M}_{K,\text{true}}\} \\ &= \mathcal{C}'_{\text{true}} \times \{\boldsymbol{U}_{1,\text{true}}, \ldots, \boldsymbol{U}_{K,\text{true}}\}, \end{aligned} \tag{A.10}$$

where $\boldsymbol{U}_{k.\text{true}} \in \mathbb{O}(p_k, r_k)$ are orthonormal matrices and $\mathcal{C}'_{\text{true}} \in \mathbb{R}^{r\times\cdots\times r}$ is a full rank core tensor. By definition of quantities in (A.1), we have

$$\lambda' := \min_{k\in[K]} \sigma_{\min}(\text{Unfold}_k(\mathcal{S}_{\text{true}})) \in [\lambda\underline{\gamma},\ \lambda\bar{\gamma}]. \tag{A.11}$$

Now our setup shares the same setting as in Zhang and Xia (2018, Theorem 1). We summarize the relationships between our algorithm outputs and the ones in Zhang and Xia (2018). For all $k \in [K]$,

1. $\boldsymbol{M}_{k,\text{true}} = \text{SVD}_{r_k}\left(\boldsymbol{R}_k^{-1}\boldsymbol{U}_{k,\text{true}}\right) :=$ the first r_k left singular vectors of $\boldsymbol{R}_k^{-1}\boldsymbol{U}_{k,\text{true}}$;
2. $\hat{\boldsymbol{M}}_k^{(t)} = \text{SVD}_{r_k}\left(\boldsymbol{R}_k^{-1}\hat{\boldsymbol{U}}_k^{(t)}\right)$ for all $t = 0, 1, 2, \ldots$;

where $\hat{\boldsymbol{U}}_k^{(t)}$ denotes the t-th iteration output of Higher Order Orthogonal Iteration (HOOI) algorithm (Zhang and Xia, 2018) with inputs $\bar{\mathcal{Y}}$. The first relationship is from (A.10), and second relationship is from induction by t. Briefly, $t = 0$ holds because of the definition $\hat{\boldsymbol{M}}_k^{(0)}$ based on lines 4-5 of our initialization Algorithm 2. For $t \geq 1, \ldots$, notice that $\hat{\boldsymbol{M}}_k^{(t)}$ is an optimizer of the objective

$$\|\bar{\mathcal{Y}} - \hat{\mathcal{C}}^{(t-1)} \times \{\boldsymbol{R}_1\hat{\boldsymbol{M}}_1^{(t)}, \ldots, \boldsymbol{R}_{k-1}\hat{\boldsymbol{M}}_{k-1}^{(t)}, \boldsymbol{R}_k\boldsymbol{M}, \boldsymbol{R}_{k+1}\hat{\boldsymbol{M}}_{k+1}^{(t-1)}, \ldots, \boldsymbol{R}_K\hat{\boldsymbol{M}}_K^{(t-1)}\}\|_F^2,$$

from the line 3 of Algorithm 1. By unfolding along the mode k, the optimizer $\boldsymbol{M}_k^{(t)}$ must satisfy

$$\begin{aligned}&\text{Unfold}_k\left(\bar{\mathcal{Y}} \times \left\{(\hat{\boldsymbol{M}}_1^{(t)})^T\boldsymbol{R}_1^{-1}, \ldots, (\hat{\boldsymbol{M}}_{k-1}^{(t)})^T\boldsymbol{R}_{k-1}^{-1},\ \boldsymbol{I}_{p_k},\ (\hat{\boldsymbol{M}}_{k+1}^{(t-1)})^T\boldsymbol{R}_{k+1}^{-1}, \ldots, (\hat{\boldsymbol{M}}_K^{(t-1)})^T\boldsymbol{R}_K^{-1}\right\}\right)\\ &\quad = \boldsymbol{R}_k\hat{\boldsymbol{M}}_k^{(t)}\text{Unfold}_k\left(\hat{\mathcal{C}}^{(t-1)}\right)\left(\boldsymbol{I}_{r_K} \otimes \cdots \otimes \boldsymbol{I}_{r_{k+1}} \otimes \boldsymbol{I}_{r_{k-1}} \otimes \boldsymbol{I}_{r_1}\right).\end{aligned} \tag{A.12}$$

Notice that the first r_k left singular vectors of the left side of (A.12) is $\hat{\boldsymbol{U}}_k^{(t)}$ in HOOI algorithm. Therefore, we prove the the second relationship by induction.

Combination of Lemma A.5 and the relationships between our algorithm out-

puts and the ones in Zhang and Xia (2018) gives us

$$\left(\frac{\underline{\gamma}}{\bar{\gamma}}\right)^2 \max_{k\in[K]} \sin\Theta(\boldsymbol{U}_{k,\text{true}}, \hat{\boldsymbol{U}}_k^{(t)}) \le \max_{k\in[K]} \sin\Theta(\boldsymbol{M}_{k,\text{true}}, \hat{\boldsymbol{M}}_k^{(t)}) \le \left(\frac{\bar{\gamma}}{\underline{\gamma}}\right)^2 \max_{k\in[K]} \sin\Theta(\boldsymbol{U}_{k,\text{true}}, \hat{\boldsymbol{U}}_k^{(t)}). \tag{A.13}$$

Now, we prove the property (a) in Proposition 2.4. Based on Lemma A.4(a), whenever $\lambda'/\sqrt{\phi} \ge C_{\text{gap}} p^{K/4}$, we have

$$\max_{k\in[K]} \sin\Theta(\boldsymbol{U}_{k,\text{true}}, \hat{\boldsymbol{U}}_k^{(0)}) \le c\left(\frac{p^{K/2}}{(\lambda\underline{\gamma})^2/\phi}\right), \tag{A.14}$$

with probability at least $1-\exp(-p)$. Notice that

$$\lambda' \overset{\text{(A.11)}}{\ge} \lambda\underline{\gamma} \gtrsim \lambda d^{K/2} \ge C_{\text{gap}}\sqrt{\phi} p^{K/4},$$

where the second inequality uses [Case 1] and [Case 2] in the proof of Theorem 2.2. The condition $\lambda/\sqrt{\phi} \ge C p^{K/4} d^{-K/2}$ guarantees a sufficiently large C_{gap} that satisfies $\lambda'/\sqrt{\phi} \ge C_{\text{gap}} p^{K/4}$. Thus combining (A.13) and (A.14) yields

$$\begin{aligned}\max_{k\in[K]} \sin\Theta(\boldsymbol{M}_{k,\text{true}}, \hat{\boldsymbol{M}}_k^{(0)}) &\le \left(\frac{\bar{\gamma}}{\underline{\gamma}}\right)^2 \left(\frac{\sqrt{\phi} p^{K/4}}{\lambda\underline{\gamma}}\right)^2 \\ &\le \frac{1}{2},\end{aligned}$$

where the last inequality uses the fact that $\underline{\gamma} \asymp d^{K/2}$ and $\bar{\gamma}/\underline{\gamma}$ is bounded by a constant in [Case 1] and [Case 2], and the condition $\lambda/\sqrt{\phi} \ge C p^{K/4} d^{-K/2}$.

Now, we prove the property (b) in Proposition 2.4. Based on Lemma A.4(b), we have

$$\max_{k\in[K]} \sin\Theta(\boldsymbol{U}_{k,\text{true}}, \hat{\boldsymbol{U}}_k^{(t)}) \lesssim \frac{\sqrt{p\phi}}{\lambda\underline{\gamma}} + \left(\frac{1}{2}\right)^t \max_{k\in[K]} \sin\Theta(\boldsymbol{U}_{k,\text{true}}, \hat{\boldsymbol{U}}_k^{(0)}),$$

with probability at least $1-\exp(-p)$. Combining (A.13) with the above inequality

yields

$$\begin{aligned}\max_{k\in[K]} \sin\Theta(\boldsymbol{M}_{k,\text{true}}, \hat{\boldsymbol{M}}_k^{(t)}) &\lesssim \max_{k\in[K]} \sin\Theta(\boldsymbol{U}_{k,\text{true}}, \hat{\boldsymbol{U}}_k^{(t)}) \\ &\lesssim \frac{\sqrt{p\phi}}{\lambda\underline{\gamma}} + \left(\frac{1}{2}\right)^t \max_{k\in[K]} \sin\Theta(\boldsymbol{U}_{k,\text{true}}, \hat{\boldsymbol{U}}_k^{(0)}) \\ &\lesssim \frac{\sqrt{p\phi}}{\lambda\underline{\gamma}} + \left(\frac{1}{2}\right)^t \max_{k\in[K]} \sin\Theta(\boldsymbol{M}_{k,\text{true}}, \hat{\boldsymbol{M}}_k^{(0)}).\end{aligned}$$

Finally, the proof is completed applying $\underline{\gamma} \asymp d^{K/2}$ from [Case 1] and [Case 2]. □

Proof of Theorem 2.3. Combining Proposition 2.4(b) and (A.14), we obtain

$$\max_{k\in[K]} \sin\Theta(\boldsymbol{U}_{k,\text{true}}, \hat{\boldsymbol{U}}_k^{(t)}) \lesssim \frac{\sqrt{p\phi}}{\lambda\underline{\gamma}} + \left(\frac{1}{2}\right)^t \left(\frac{p^{K/2}}{(\lambda\underline{\gamma})^2/\phi}\right),$$

with probability at least $1 - \exp(-p)$. We set $t \gtrsim \log \frac{p^{(K-1)/2}}{\lambda\underline{\gamma}}$ to make the second term negligible. Therefore, the first part of proof is completed by noticing that

$$\frac{p^{(K-1)/2}}{\lambda\underline{\gamma}} \lesssim \log \frac{p^{(K-1)/2}}{\lambda d^{K/2}} \lesssim K \log p,$$

where the first inequality uses $\underline{\gamma} \asymp d^{K/2}$ from [Case 1] and [Case 2], and the last inequality is from the condition $\lambda/\sqrt{\phi} \geq Cp^{K/4}d^{-K/2}$.

For the estimation error with respect to Frobenius norm, direct application of Lemma A.4(c) with $t \gtrsim K \log p \gtrsim \log \frac{p^{(K-1)/2}}{\lambda\underline{\gamma}}$ yields

$$\|\hat{\mathcal{S}}^{(t)} - \mathcal{S}_{\text{true}}\|_F^2 \lesssim \phi(r^K + Kpr), \tag{A.15}$$

with probability at least $1 - \exp(-p)$. Notice that

$$\begin{aligned}\|\hat{\mathcal{S}}^{(t)} - \mathcal{S}_{\text{true}}\|_F^2 &= \|\left(\hat{\mathcal{B}}^{(t)} - \mathcal{B}_{\text{true}}\right) \times \{\boldsymbol{R}_1, \dots \boldsymbol{R}_K\}\|_F^2 \\ &\geq \underline{\gamma}^2 \|\hat{\mathcal{B}}^{(t)} - \mathcal{B}_{\text{true}}\|_F^2 \\ &\gtrsim d^K \|\hat{\mathcal{B}}^{(t)} - \mathcal{B}_{\text{true}}\|_F^2, \quad \text{from [Case 1] and [Case 2].}\end{aligned} \tag{A.16}$$

Combining (A.15) and (A.16) completes the proof. □

Auxiliary Lemmas

Proof of Proposition A.1. For ease of presentation, we drop the subscript $(i_1, \ldots, i_K)$ and simply write ε $(= y - b'(\theta))$. For any given $t \in \mathbb{R}$, we have

$$\begin{aligned}
\mathbb{E}(\exp(t\varepsilon|\theta)) &= \int c(x) \exp\left(\frac{\theta x - b(\theta)}{\phi}\right) \exp\left(t(x - b'(\theta))\right) dx \\
&= \int c(x) \exp\left(\frac{(\theta + \phi t)x - b(\theta + \phi t) + b(\theta + \phi t) - b(\theta) - \phi t b'(\theta)}{\phi}\right) dx \\
&= \exp\left(\frac{b(\theta + \phi t) - b(\theta) - \phi t b'(\theta)}{\phi}\right) \\
&\leq \exp\left(\frac{\phi U t^2}{2}\right),
\end{aligned}$$

where $c(\cdot)$ and $b(\cdot)$ are known functions in the exponential family corresponding to y, and the last line uses the fact that $\sup_{\theta \in \mathbb{R}} b''(\theta) \leq U$. Therefore, ε is sub-Gaussian-(ϕU). □

Definition A.1 (α-convexity)**.** *A real-valued function $f \colon \mathcal{S} \to \mathbb{R}$ is called α-convex, if*

$$f(x_1) \geq f(x_2) + \langle \nabla_x f(x_2),\ x_1 - x_2 \rangle + \alpha \|x_1 - x_2\|_F^2, \text{ for all } x_1, x_2 \in \mathcal{S}.$$

Lemma A.2 (Convexity under linear transformation)**.** *Suppose $f \colon \mathbb{R}^{d \times \cdots \times d} \to \mathbb{R}$ is a α-convex function. Define a function $g \colon \mathbb{R}^{p \times \cdots \times p} \to \mathbb{R}$ by $g(\mathcal{B}) = f(\mathcal{B} \times \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K\})$ for all $\mathcal{B} \in \mathbb{R}^{p \times \cdots \times p}$. Then, g is a $(\underline{\gamma}^2 \alpha)$-convex function.*

Proof of Lemma A.2. By the definition of α-convexity, we have

$$f(\Theta_1) \geq f(\Theta_2) + \langle \nabla_\Theta f(\Theta_2), \Theta_1 - \Theta_2 \rangle + \alpha \|\Theta_1 - \Theta_2\|_F^2, \text{ for all } \Theta_1, \Theta_2 \in \mathbb{R}^{d \times \cdots \times d}, \tag{A.17}$$

where $\nabla_\Theta f(\cdot)$ denotes the derivative of f with respect to $\Theta \in \mathbb{R}^{d \times \cdots \times d}$. For any $\mathcal{B}_1, \mathcal{B}_2 \in \mathbb{R}^{p \times \cdots \times p}$, we notice that $\mathcal{B}_i \times \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K\} \in \mathbb{R}^{d \times \cdots \times d}$ for $i = 1, 2$. Apply-

ing (A.17) to this setting gives

$$
\begin{aligned}
& f(\mathcal{B}_1 \times \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K\}) \\
\geq\ & f(\mathcal{B}_2 \times \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K\}) + \langle \nabla_\Theta f(\mathcal{B}_2 \times \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K\}), (\mathcal{B}_1 - \mathcal{B}_2) \times \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K\} \rangle \\
& + \alpha \|(\mathcal{B}_1 - \mathcal{B}_2) \times \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K\}\|_F^2 \\
\geq\ & f(\mathcal{B}_2 \times \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K\}) + \langle \nabla_\Theta f(\mathcal{B}_2 \times \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K\}) \times \{\boldsymbol{X}_1^T, \ldots, \boldsymbol{X}_K^T\}, (\mathcal{B}_1 - \mathcal{B}_2) \rangle \\
& + \alpha \underline{\gamma}^2 \|\mathcal{B}_1 - \mathcal{B}_2\|_F^2. \qquad \text{(A.18)}
\end{aligned}
$$

By the definition of g and the linearity from $\mathcal{B}$ to Θ, we have

$$
\nabla g_{\mathcal{B}}(\mathcal{B}_2) = \nabla f_\Theta(\mathcal{B}_2 \times \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K\}) \times \{\boldsymbol{X}_1^T, \ldots, \boldsymbol{X}_K^T\}. \qquad \text{(A.19)}
$$

The convexity of g directly follows by plugging (A.19) into (A.18),

$$
g(\mathcal{B}_1) \geq g(\mathcal{B}_2) + \langle \nabla g_{\mathcal{B}}(\mathcal{B}_2), \mathcal{B}_1 - \mathcal{B}_2 \rangle + \alpha \underline{\gamma}^2 \|\mathcal{B}_1 - \mathcal{B}_2\|_F^2.
$$

□

Proof of Proposition A.2. We first prove the strong concavity by viewing the log-likelihood as a function of the linear predictor Θ. Write

$$
\bar{\mathcal{L}}(\Theta) = \langle \mathcal{Y}, \Theta \rangle - \sum_{i_1, \ldots, i_K} b(\theta_{i_1, \ldots, i_K}).
$$

Direct calculation shows that the Hession of $\bar{\mathcal{L}}(\Theta)$ can be expressed as

$$
\frac{\partial \bar{\mathcal{L}}(\Theta)}{\partial \theta_{i_1, \ldots, i_K} \partial \theta_{j_1, \ldots, j_K}} = \begin{cases} -b''(\theta_{i_1, \ldots, i_K}) < -L < 0, & \text{if } (i_1, \ldots, i_K) = (j_1, \ldots, j_K), \\ 0, & \text{otherwise}, \end{cases}
$$

Therefore, the Hession matrix of $\bar{\mathcal{L}}(\Theta)$ is strictly negative definite with eigenvalues upper bounded by $-L < 0$. By Taylor expansion, $-\bar{\mathcal{L}}(\Theta)$ is $L/2$-convex with respect to Θ. Note that $\bar{\mathcal{L}}(\Theta) = \mathcal{L}(\mathcal{B})$ via the linear mapping $\Theta = \mathcal{B} \times \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K\}$. Therefore, by Lemma A.2, $\mathcal{L}(\mathcal{B})$ is $(\gamma^2 L/2)$-convex with respect to $\mathcal{B}$.

To prove the second part of Proposition A.2, we note

$$\langle \nabla \mathcal{L}(\mathcal{B}_{\text{true}}), \mathcal{B} \rangle = \langle \nabla \bar{\mathcal{L}}(\Theta_{\text{true}}) \times \{\boldsymbol{X}_1^T, \ldots, \boldsymbol{X}_K^T\},\, \mathcal{B} \rangle = \langle \mathcal{Y} - b'(\Theta_{\text{true}}),\, \mathcal{B} \times \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K\} \rangle.$$

By Proposition A.1, $\mathcal{Y} - b'(\Theta_{\text{true}})$ is a random tensor consisting of i.i.d. sub-Gaussian-$(U\phi)$ entries under Assumption 2. We write $\mathcal{E} = \mathcal{Y} - b'(\Theta_{\text{true}})$ and consider the sub-Gaussian maxima

$$\text{Err}_{\text{ideal}}(\boldsymbol{r}) = \sup_{\|\mathcal{B}\|_F = 1, \mathcal{B} \in \mathcal{P}(r)} \langle \mathcal{E}, \mathcal{B} \times \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K\} \rangle.$$

The quantity $\text{Err}_{\text{ideal}}(\boldsymbol{r})$ is closely related to the localized Gaussian width (Chen et al., 2019a; Han et al., 2020) that measures the model complexity of $\mathcal{P}(\boldsymbol{r})$. By adapting Han et al. (2020, Lemma E.5) in our context, we have

$$\text{Err}_{\text{ideal}}(\boldsymbol{r}) \lesssim \sqrt{\phi U (r^K + Kpr)} \prod_{k \in [K]} \sigma_{\max}(\boldsymbol{X}_k) \leq \bar{\gamma} \sqrt{\phi U (r^K + Kpr)},$$

with probability at least $1 - \exp(-p)$. □

The following Lemma is adopted from Wang and Song (2017, Theorem 6.1) in our contexts.

Lemma A.3 (Wedin's $\sin\Theta$ Theorem)**.** *Let $\boldsymbol{B}$ and $\hat{\boldsymbol{B}}$ be two $m \times n$ real matrix SVDs $\boldsymbol{B} = \boldsymbol{U}\Sigma\boldsymbol{V}^T$ and $\hat{\boldsymbol{B}} = \hat{\boldsymbol{U}}\hat{\boldsymbol{\Sigma}}\hat{\boldsymbol{V}}^T$. If $\sigma_{min}(\boldsymbol{B}) > 0$ and $\|\hat{\boldsymbol{B}} - \boldsymbol{B}\|_F \ll \sigma_{min}(\boldsymbol{B})$, then*

$$sin\Theta(\boldsymbol{U}, \hat{\boldsymbol{U}}) \leq \frac{\sigma_{\max}(\hat{\boldsymbol{B}} - \boldsymbol{B})}{\sigma_{min}(\boldsymbol{B})} \leq \frac{\|\hat{\boldsymbol{B}} - \boldsymbol{B}\|_F}{\sigma_{min}(\boldsymbol{B})}.$$

The following theorem Zhang and Xia (2018) provides the statistical guarantees for unsupervised tensor decomposition based on alternating least square algorithm. For simplicity, we consider the balanced dimension $p_1 = \cdots = p_K = p$ and $r_1 = \cdots = r_K = r$.

Lemma A.4 (Theorem 1 in Zhang and Xia (2018))**.** *Consider the Gaussian tensor model*

$$\mathcal{Y} = \mathcal{S}_{true} + \mathcal{E},$$

where $\mathcal{S}_{true} = \mathcal{C}_{true} \times \{\boldsymbol{U}_{1,true}, \ldots, \boldsymbol{U}_{K,true}\}$ *is an unknown signal tensor,* $\mathcal{C}_{true} \in \mathbb{R}^{r \times \cdots \times r}$ *is a full rank core tensor,* $\boldsymbol{U}_{k,true} \in \mathbb{O}(p, r)$ *are orthornomal matrices, and* $\mathcal{E} \in \mathbb{R}^{p \times \cdots \times p}$ *is a Gaussian noise tensor consisting of i.i.d entries from* $N(0, \sigma)$*. Let* λ *denote the smallest singular value of matrices* $Unfold_k(\mathcal{S}_{true})$ *over all possible* k*,*

$$\lambda' = \min_{k \in [K]} \sigma_{\min}(Unfold_k(\mathcal{S}_{true})).$$

Then, the following two properties hold whenever $\lambda'/\sigma \geq C_{\text{gap}} p^{K/4}$ *for some universal constant* $C_{\text{gap}} > 0$*.*

(*a*) *With probability at least* $1 - \exp(-p)$*, the spectral initialization* $\hat{\boldsymbol{U}}_k^{(0)}$ *has*

$$\max_{k \in [K]} \sin \Theta(\boldsymbol{U}_{k,true}, \hat{\boldsymbol{U}}_k^{(0)}) \leq c \frac{p^{K/2}}{\lambda'^2/\sigma^2},$$

for some constant $c > 0$*.*

(*b*) *Let* $t = 1, 2, \ldots,$ *denote the iteration in HOOI algorithm. With probability at least* $1 - \exp(-p)$*, the alternating optimization* $\hat{\boldsymbol{U}}_k^{(t)}$ *satisfies*

$$\max_{k \in [K]} \sin \Theta(\boldsymbol{U}_{k,true}, \hat{\boldsymbol{U}}_k^{(t)}) \lesssim \frac{\sqrt{p}}{\lambda'/\sigma} + \left(\frac{1}{2}\right)^t \max_{k \in [K]} \sin \Theta(\boldsymbol{U}_{k,true}, \hat{\boldsymbol{U}}_k^{(0)}),$$

(*c*) *When* $t \gtrsim \log \frac{p^{(K-1)/2}}{\lambda'}$*, the tensor estimate* $\hat{\mathcal{S}}^{(t)}$ *from HOOI satisfies*

$$\|\hat{\mathcal{S}}^{(t)} - \mathcal{S}_{true}\|_F^2 \lesssim \sigma^2(r^K + Kpr),$$

with probability at least $1 - \exp(-p)$*.*

Lemma A.5 (Angle distance under linear transformation)**.** *Let $\boldsymbol{U}$ and $\hat{\boldsymbol{U}}$ be two $m \times n$ real matrices where $m > n$. Let $\boldsymbol{R}$ be an $m \times m$ invertible matrix. If $\sin\Theta(\boldsymbol{U}, \hat{\boldsymbol{U}}) \leq L$ for some constant $L \in [0, 1]$, then*

$$\sin\Theta(\boldsymbol{RU}, \boldsymbol{R}\hat{\boldsymbol{U}}) \leq \left(\frac{\sigma_{\max}(\boldsymbol{R})}{\sigma_{\min}(\boldsymbol{R})}\right)^2 L.$$

Proof. Suppose that orthonormal basis of Span($\boldsymbol{U}$) and Span($\hat{\boldsymbol{U}}^{\perp}$) are $\{\mu_1, \ldots, \mu_n\}$ and $\{\nu_{n+1}, \ldots, \nu_m\}$ respectively. By definition,

$$\sin\Theta(\boldsymbol{U}, \hat{\boldsymbol{U}}) = \max_{\sum_{i=1}^n a_i^2 = \sum_{j=n+1}^m b_j^2 = 1} \left\langle \sum_{i=1}^n a_i \mu_i, \sum_{j=n+1}^m b_j \nu_j \right\rangle \leq L.$$

We write $\boldsymbol{x} = \boldsymbol{R}\sum_{i=1}^n a_i\mu_i$ and $\boldsymbol{y} = \boldsymbol{R}\sum_{j=n+1}^m b_j\nu_j$ for any $\boldsymbol{x} \in$ Span($\boldsymbol{RU}$) and $\boldsymbol{y} \in$ Span($(\boldsymbol{R}\hat{\boldsymbol{U}})^{\perp}$). Then,

$$\begin{aligned}
\frac{\langle \boldsymbol{x}, \boldsymbol{y}\rangle}{\|\boldsymbol{x}\|_2\|\boldsymbol{y}\|_2} &= \frac{\langle \boldsymbol{R}\sum_{i=1}^n a_i\mu_i, \boldsymbol{R}\sum_{j=n+1}^m b_j\nu_j\rangle}{\|\boldsymbol{R}\sum_{i=1}^n a_i\mu_i\|_2\|\boldsymbol{R}\sum_{j=n+1}^m b_j\nu_j\|_2} \\
&\leq \frac{\sigma_{\max}(\boldsymbol{R}^T\boldsymbol{R})\langle\sum_{i=1}^n a_i\mu_i, \sum_{j=n+1}^m b_j\nu_j\rangle}{\sigma_{\min}^2(\boldsymbol{R})\sqrt{\sum_{i=1}^n a_i^2}\sqrt{\sum_{j=n+1}^m b_j^2}} \\
&\leq \left(\frac{\sigma_{\max}(\boldsymbol{R})}{\sigma_{\min}(\boldsymbol{R})}\right)^2 \sin\Theta(\boldsymbol{U}, \hat{\boldsymbol{U}}).
\end{aligned}$$

□

A.2 Additional simulation results

Detailed simulation setup for Figure 2.6a-b

We generate data from **Envelope** model (Li and Zhang, 2017) with slight modification. We simulate response tensor $\mathcal{Y} \in \mathbb{R}^{d\times d\times d}$ from the following model with

envelope dimension (u_1, u_2),

$$\begin{aligned} \mathcal{Y}|\boldsymbol{X} &= \mathcal{B} \times_3 \boldsymbol{X} + \mathcal{E} = \mathcal{C} \times \{\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2, \boldsymbol{X}\} + \mathcal{E}, \\ \text{with} \quad \mathcal{E} &\sim \mathcal{TN}(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \boldsymbol{I}), \boldsymbol{\Sigma}_k = \boldsymbol{\Gamma}_k \boldsymbol{\Omega}_k \boldsymbol{\Gamma}_k^T + \boldsymbol{\Gamma}_{0k} \boldsymbol{\Omega}_{0k} \boldsymbol{\Gamma}_{0k}^T + \boldsymbol{I}, k = 1, 2, \end{aligned} \tag{A.20}$$

where $\boldsymbol{X} \in \mathbb{R}^{d \times p}$ is the feature matrix, $\mathcal{B} = \mathcal{C} \times \{\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2, \boldsymbol{I}\} \in \mathbb{R}^{d \times d \times p}$ is the coefficient tensor, $\mathcal{C} \in \mathbb{R}^{\mu_1 \times \mu_2 \times p}$ is a full-rank core tensor, $\mathcal{TN}(\cdot,\ \cdot,\ \cdot)$ represents zero-mean tensor normal distribution with Kronecker structured covariance, $\boldsymbol{\Gamma}_k \in \mathbb{O}(d, u_k)$ consists of orthogonal columns, $\boldsymbol{\Gamma}_{0k} \in \mathbb{O}(d, d - u_k)$ is the orthogonal complement of $\boldsymbol{\Gamma}_k$, and $\boldsymbol{\Omega}_k = \boldsymbol{A}_k \boldsymbol{A}_k^T$, $\boldsymbol{\Omega}_{0k} = \boldsymbol{A}_{k0} \boldsymbol{A}_{k0}^T$ with $\boldsymbol{A}_k \in \mathbb{R}^{u_k \times u_k}$, $\boldsymbol{A}_{k0} \in \mathbb{R}^{(d-u_k) \times (d-u_k)}$.

The entries of $\boldsymbol{X}$ are i.i.d. drawn from $\mathcal{N}(0, 1)$, the entries of $\boldsymbol{A}_k$, $\boldsymbol{A}_{k0}$ are i.i.d. drawn from Uniform$[-\gamma, \gamma]$, and the entries of core tensor $\mathcal{C}$ are i.i.d. drawn from Uniform$[-3, 3]$. We call γ the *correlation level*. Note that the only distinction between model (A.20) and standard **Envelope** model is the additional identity matrix $\boldsymbol{I}$ in the expression of $\boldsymbol{\Sigma}_k$. When $\gamma = 0$, the model (A.20) reduces to our **STD** model with rank $\boldsymbol{r} = (u_1, u_2, p)$. We set $d = 20, p = 5$ in our simulation.

Detailed simulation setup for Figure 2.6c-d

We generate the data from **GLSNet** model (Zhang et al., 2018) with slight modification. We simulate the binary response tensor $\mathcal{Y} \in \{0, 1\}^{d \times d \times d}$ from the following model

$$\mathbb{E}[\mathcal{Y}|\boldsymbol{X}] = f(\boldsymbol{1} \otimes \boldsymbol{\Theta} + \mathcal{B} \times_3 \boldsymbol{X}),$$

where $f(\cdot)$ is the logistic link, $\boldsymbol{X} \in \mathbb{O}(d, p)$ is the feature matrix with orthonormal columns, $\boldsymbol{\Theta} = \boldsymbol{A}\boldsymbol{A}^T \in \mathbb{R}^{d \times d}$ is a rank-R intercept matrix, where the entries of $\boldsymbol{A} \in \mathbb{R}^{d \times R}$ are simulated from i.i.d. standard normal. Unlike original **GLSNet** model, we generate joint sparse and low-rank structure to the coefficient tensor $\mathcal{B}$ as follows.

To generate $\mathcal{B}$, we firstly generate a low-rank tensor $\mathcal{B}_0$ as

$$\mathcal{B}_0 = \mathcal{C} \times \boldsymbol{M}_1 \times \boldsymbol{M}_2 \times \boldsymbol{M}_3,$$

where $\mathcal{C} \in \mathbb{R}^{R \times R \times R}$ is a full-rank core tensor, $\boldsymbol{M}_1, \boldsymbol{M}_2 \in \mathbb{R}^{d \times R}$ and $\boldsymbol{M}_3 \in \mathbb{R}^{p \times R}$ are the factor matrices with orthonormal columns. We simulate i.i.d. uniform entries in $\mathcal{C}$ and rescale the tensor $\mathcal{B}_0$ such that $\|\mathcal{B}_0\|_{\max} = 2$. Last, we obtain a sparse $\mathcal{B}$ by randomly setting sd^2p entries in $\mathcal{B}_0$ to zero. We call s the *sparsity level* which quantifies the proportion of zero's in $\mathcal{B}$. Hence, the generated tensor $\mathcal{B}$ is of sparsity level s and of low-rank (R, R, R). We set $d = 20$, $p = 5$ and consider the combination of rank $R = 2$ (low), 4 (high) and sparsity level $s = \{0, 0.3, 0.5\}$ in the simulation.

Comparison with GLMs under stochastic block models

We investigate the performance of our model under correlated feature effects. We mimic the scenario of brain imaging analysis. A sample of $d_3 = 50$ networks are simulated, one for each individual. Each network measures the connections between $d_1 = d_2 = 20$ brain nodes. We simulate $p = 5$ features for the each of the 50 individuals. These features may represent, for example, age, gender, cognitive score, etc. Recent study has suggested that brain connectivity networks often exhibit community structure represented as a collection of subnetworks, and each subnetwork is comprised of a set of spatially distributed brain nodes. To accommodate this structure, we utilize the stochastic block model (Abbe, 2017) to generate the effect size. Specifically, we partition the nodes into r blocks by assigning each node to a block with uniform probability. Edges within a same block are assumed to share the same feature effects, where the effects are i.i.d. drawn from $N(0, 1)$. We then apply our tensor regression model to the network data using the BIC-selected rank. Note that in this case, the true model rank is unknown; the rank of a r-block network is not necessarily equal to matrix rank r (Wang and Zeng, 2019).

Figure A.1 compares the MSE of our method with a multiple-response GLM approach. The multiple-response GLM is to regress the dyadic edges, one at a time,

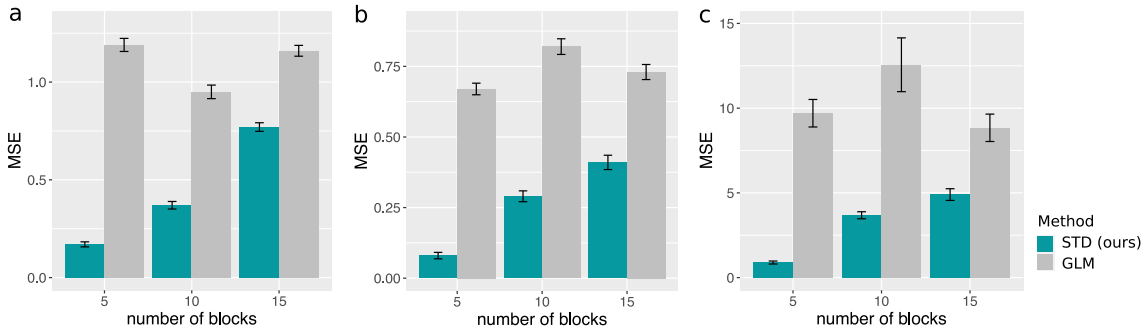

Figure A.1: Performance comparison under stochastic block models. The three panels plot the MSE when the response tensors are generated from (a) Gaussian (b) Poisson and (c) Bernoulli models. The x-axis represents the number of blocks in the networks.

on the features, and this model is repeatedly fitted for each edge. As we find in Figure A.1, our tensor regression method achieves significant error reduction in all three data types considered. The outperformance is substantial in the presence of large communities; even in the less structured case ($\sim 20/15 = 1.33$ nodes per block), our method still outer-performs GLM. The possible reason is that the multiple-response GLM approach does not account for the correlation among the edges, and suffers from overfitting. In contrast, the low-rankness in our modeling incorporates the shared information across entries. By selecting the rank in a data-driven way, our method achieves accurate estimation in a wide range of settings.

A.3 Additional results on data application

Rank selection for Nations data

Table A.1 summarizes the BIC results in the grid search $\boldsymbol{r} \in \{3, 4, 5\}^3$. We set $r_1 = r_2$ due to the symmetry in the dataset. Table A.1 shows that $(r_1, r_2) = (4, 4)$ consistently provides the minimal BIC under a range of r_3. Because multiple values of r_3 give similar BIC, we choose r_3 based on the interpretability of the results. Tables A.2-A.4

compare the clustering results for $r_3 = 3, 4, 5$. For ease of visualisation, we list only the subset of relations for which the three configurations yield incoherent clustering. We find that the clustering with $r_3 = 4$ (Table A.3) provides the cleanest results. Table A.2 with $r_3 = 3$ mixes the categories Economics with Organization and Military. Table A.4 with $r_3 = 5$ mixes Economics with Organization, while splitting Military and Territory into different clusters. Therefore, we choose the rank $\boldsymbol{r} = (4, 4, 4)$ in the main paper. The running time for the rank selection via grid search is 95 secs in total, on an iMac macOS High Sierra 10.13.6 with Intel Core i5 3.8 GHz CPU and 8 GB RAM. This indicates the BIC is feasible in the considered setting.

r_3	$r_3 = 3$			$r_3 = 4$			$r_3 = 5$		
(r_1, r_2)	$(3,3)$	$(4,4)$	$(5,5)$	$(3,3)$	$(4,4)$	$(5,5)$	$(3,3)$	$(4,4)$	$(5,5)$
BIC	11364	**11194**	11701	12275	**11897**	12365	17652	**12666**	18146

Table A.1: BIC results for *Nations* data under different tensor rank. Bold number indicates the minimal BIC with a certain r_3.

Cluster	Relations
I	exportbooks, relexportbooks, protests, tourism, reltourism, relintergovorgs relngo, intergovorgs3, ngoorgs3, militaryalliance,commonbloc1
II	militaryactions, severdiplomatic,expeldiplomats, commonbloc0, aidenemy attackembassy, lostterritory, blockpositionindex
III	tourism3, exports, relexports, exports3, intergovorgs, ngo ,embassy, reldiplomacy, commonbloc2

Economics Military Organization Territory

Table A.2: K-mean relations clustering with $r_3 = 3$. For visualization purpose, only a subset of relations are presented. See texts for details.

Cluster	Relations
I	aidenemy, attackembassy, lostterritory
II	militaryactions, severdiplomatic, expeldiplomats, protests, commonbloc0, blockpositionindex, commonbloc1
III	relintergovorgs, relngo, intergovorgs3, ngoorgs3, militaryalliance, commonbloc2
IV	exportbooks, relexportbooks, tourism, reltourism, tourism3 exports, relexports, exports3, intergovorgs, ngo, embassy, reldiplomacy

Economics Military Organization Territory

Table A.3: K-mean relations clustering with $r_3 = 4$. For visualization purpose, only a subset of relations are presented. See texts for details.

Cluster	Relations
I	exportbooks, relexportbooks, tourism, reltourism, tourism3, exports, relexports, exports3 intergovorgs, relintergovorgs, ngo, relngo, intergovorgs3, ngoorgs3, embassy, reldiplomacy
II	attackembassy
III	commonbloc0, blockpositionindex
IV	militaryalliance, commonbloc2
V	militaryactions, severdiplomatic, expeldiplomats, aidenemy, lostterritory, protests, commonbloc1

Economics Military Organization Territory

Table A.4: K-mean relations clustering with $r_3 = 5$. For visualization purpose, only a subset of relations are presented. See texts for details.

Comparison with unsupervised decomposition

We compare the supervised vs. unsupervised decomposition in the *Nations* data analysis. Table A.5 shows the clustering results based on classical unsupervised Tucker decomposition without the feature matrices. Table A.6 shows the clustering results based on supervised tensor decomposition (**STD**). Compared with supervised decomposition, the unsupervised clustering loses some interpretation.

Similar relations *exports* and *relexports*, *ngo* and *relngo* are separated into different clusters.

Cluster	Relations
I	economicaid, releconomicaid, exportbooks, relexportbooks, weightedunvote, unweightedunvote, tourism, reltourism, tourism3, exports, intergovorgs, ngo, militaryalliance
II	warning, violentactions, militaryactions, duration, severdiplomatic, expeldiplomats, boycottembargo, aidenemy, negativecomm, accusation, protests, unoffialacts, attackembassy, relemigrants, timesincewar, lostterritory, dependent
III	timesinceally, independence, commonbloc0, blockpositionindex
IV	treaties, reltreaties, officialvisits, conferences, booktranslations, relbooktranslations negativebehavior, nonviolentbehavior, emigrants, emigrants3, students, relstudents, relexports, exports3 relintergovorgs, relngo, intergovorgs3, ngoorgs3, embassy, reldiplomacy, commonbloc1, commonbloc2

Economics Military Organization Territory

Table A.5: Clustering of relations based on unsupervised tensor decomposition.

Category	Relations
I	warning, violentactions, militaryactions, duration, negativebehavior, protests, severdiplomatic timesincewar, commonbloc0, commonbloc1, blockpositionindex, expeldiplomats
II	emigrants, emigrants3, relemigrants, accusation, nonviolentbehavior, ngoorgs3, commonbloc2, intergovorgs3 releconomicaid, relintergovorgs, relngo, students, relstudents, economicaid, negativecomm, militaryalliance
III	treaties, reltreaties, officialvisits, exportbooks, relexportbooks, booktranslations, relbooktranslations boycottembargo, weightedunvote, unweightedunvote, reltourism, tourism, tourism3, exports, exports3 relexports, intergovorgs, ngo, embassy, reldiplomacy, timesinceally, independence, conferences, dependent
IV	aidenemy, lostterritory, unoffialacts, attackembassy

Economics Military Organization Territory

Table A.6: Clustering of relations based on supervised tensor decomposition.

How different are supervised vs. unsupervised factors in general?

It is helpful to realize that the unsupervised and methods address different aspects of the problem. The unsupervised decomposition identifies factors that explain most variation in the tensor, whereas the supervised decomposition identifies factors that are most attributable to side features.

We provide a simple example here for illustration.

Example A.3. *Consider the following data tensor $\mathcal{Y}$ and one-sided feature matrix $\boldsymbol{X}$,*

$$\mathcal{Y} = \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + 10\boldsymbol{e}_2 \otimes \boldsymbol{e}_2 \otimes \boldsymbol{e}_2, \quad \boldsymbol{X} = \boldsymbol{e}_1,$$

where $\boldsymbol{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ is the ith canonical basis vector in $\mathbb{R}^d$ for $i = 1, 2$. Now, consider the unsupervised vs. supervised decomposition of $\mathcal{Y}$ with rank $\boldsymbol{r} = (1, 1, 1)$. Then, the top supervised and unsupervised factors are perpendicular to each other,

$$\boldsymbol{M}_{sup,k} \perp \boldsymbol{M}_{unsup,k}, \quad \textit{for all } k = 1, 2, 3,$$

where $\boldsymbol{M}_{sup,k}$, $\boldsymbol{M}_{unsup,k}$ denote the mode-k factors from supervised and unsupervised decompositions, respectively.

Remark A.6. This example shows complementary information between factors from supervised vs. unsupervised decompositions. In general, one could construct examples such that these two methods return **arbitrarily different** factors.

B APPENDIX OF CHAPTER 3

B.1 Additional numerical experiments

Bernoulli phase transition. The first additional experiment verifies the statistical-computational gap in Section 3.3 under the Bernoulli model. Consider the Bernoulli model with $p = \{80, 100\}$, $r = 5$. We vary γ in $[-1.2, -0.4]$ and $[-2.1, -1.4]$ for matrix $(K = 2)$ and tensor $(K = 3)$ clustering, respectively. We approximate MLE using an oracle estimator, i.e., the output of Sub-algorithm 2 initialized from the true assignment. Figure B.1 shows a similar pattern as Figure 3.4. The algorithm and oracle estimators have no gap in the matrix case, while an error gap emerges between the critical values $\gamma_{\text{stat}} = -2$ and $\gamma_{\text{comp}} = -1.5$ in the tensor case. Figure 3.4 suggests the statistical-computational gap in Bernoulli models.

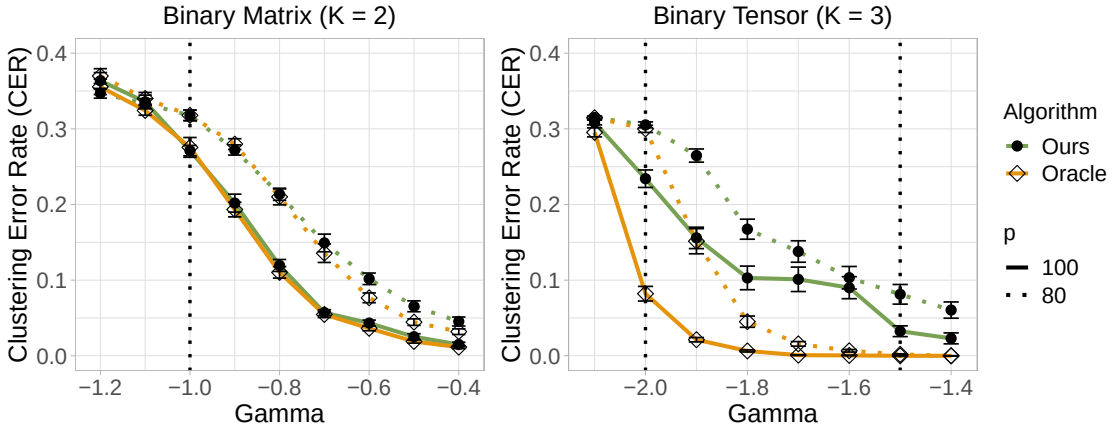

Figure B.1: SNR phase transitions for Bernoulli dTBM with $p = \{80, 100\}, r = 5$ under (a) matrix case with $\gamma \in [-1.2, -0.4]$ and (b) tensor case with $\gamma \in [-2.1, -1.4]$.

Sparsity. The second additional experiment evaluates the algorithm performances under the sparse binary dTBM (3.18). We fix the signal exponent $\gamma = -1.2$ and vary the sparsity parameter $\alpha_p \in [0.05, 0.9]$. A smaller α_p leads to a higher

probability of zero entries in the observation. In addition to the three algorithms mentioned in Section 3.6 (denoted **Initialization**, **dTBM**, and **SCORE**), we consider other three algorithms based on the discussion in Section 3.4:

- **D-HOSVD**, the diagonal-deleted HOSVD in Ke et al. (2019);
- **D-HOSVD + Angle**, the combined algorithm of our angle-based iteration with initialization from **D-HOSVD**;
- **SCORE + Angle**, the combined algorithms of our angle-based iteration with initialization from **SCORE**.

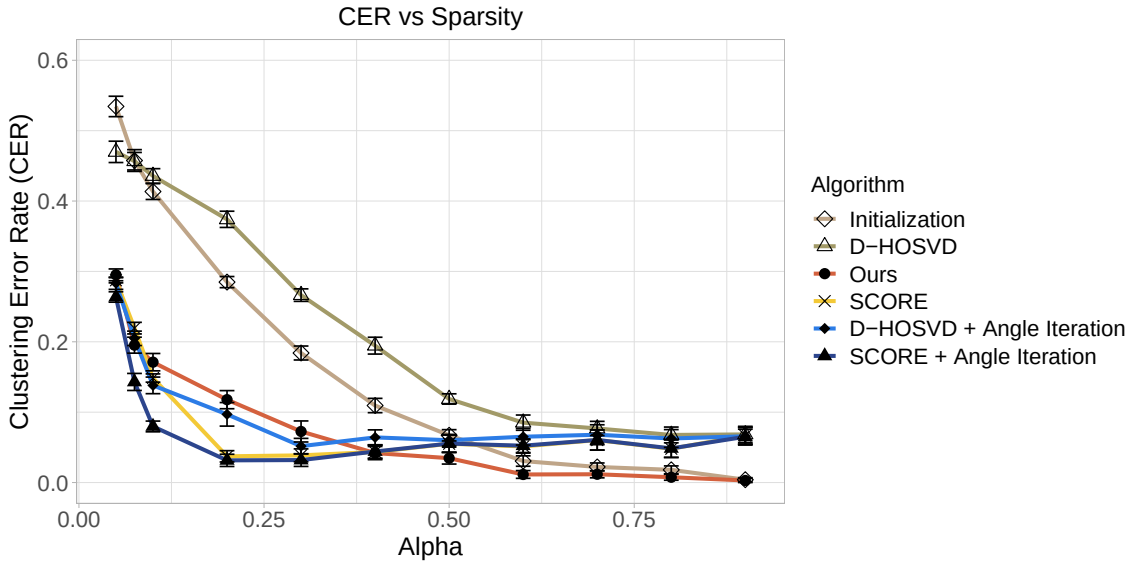

Figure B.2: CER comparison versus sparsity parameter α_p in $[0.05, 0.9]$. We set $p = 100, r = 5$ and $\gamma = -1.2$ under sparse binary dTBM.

Figure B.2 shows a slightly larger error in **dTBM** than that in **SCORE**, **D-HOSVD + Angle**, and **SCORE + Angle** under the sparse setting with $\alpha_p < 0.3$. The small gap between **dTBM** and other sparse-specific methods implies the robustness of our algorithm. In addition, comparing **SCORE** versus **SCORE + Angle** (or **D-HOSVD**

versus **D-HOSVD + Angle**) indicates the benefit of our angle iterations under the sparse dTBM. In the intermediate and dense cases with $\alpha_p \geq 0.3$, our proposed **dTBM** has a clear improvement over others, which again verifies the success of our algorithm in dense settings.

B.2 Proofs

We provide the proofs for all the theorems in our main paper. In each sub-section, we first show the proof of main theorem and then collect the useful lemmas in the end. We combine the proofs of MLE achievement in Theorem 3.7 and polynomial-time achievement in Theorem 3.17 in the last section due to the similar idea.

Notation

Before the proofs, we first introduce the notation used throughout the appendix and the general dTBM without symmetric assumptions. The parameter space and minimal gap assumption are also extended for the general asymmetric dTBM.

Preliminaries.

1. For mode $k \in [K]$, denote mode-k tensor matricizations by

$$\begin{aligned} \boldsymbol{Y}_k &= \mathrm{Mat}_k(\mathcal{Y}), \quad \boldsymbol{S}_k = \mathrm{Mat}_k(\mathcal{S}), \\ \boldsymbol{E}_k &= \mathrm{Mat}_k(\mathcal{E}), \quad \boldsymbol{X}_k = \mathrm{Mat}_k(\mathcal{X}). \end{aligned}$$

2. For a vector $\boldsymbol{a}$, let $\boldsymbol{a}^s := \boldsymbol{a}/\|\boldsymbol{a}\|$ denote the normalized vector. We make the convention that $\boldsymbol{a}^s = \boldsymbol{0}$ if $\boldsymbol{a} = \boldsymbol{0}$.

3. For a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times m}$, let $\boldsymbol{A}^{\otimes K} := \boldsymbol{A} \otimes \cdots \otimes \boldsymbol{A} \in \mathbb{R}^{n^K \times m^K}$ denote the Kronecker product of K copies of matrices $\boldsymbol{A}$.

4. For a matrix $\boldsymbol{A}$, let $\|\boldsymbol{A}\|_\sigma$ denote the spectral norm of matrix $\boldsymbol{A}$, which is equal to the maximal singular value of $\boldsymbol{A}$; let $\lambda_k(\boldsymbol{A})$ denote the k-th largest singular value of $\boldsymbol{A}$; let $\|\boldsymbol{A}\|_F$ denote the Frobenius norm of matrix $\boldsymbol{A}$.

Extension to general asymmetric dTBM.

The general order-K $(p_1, \ldots, p_K)$-dimensional dTBM with r_k communities and degree heterogeneity $\boldsymbol{\theta}_k = [\![\theta_k(i)]\!] \in \mathbb{R}_+^{p_k}$ is represented by

$$\mathcal{Y} = \mathcal{X} + \mathcal{E}, \text{ where } \mathcal{X} = \mathcal{S} \times_1 \boldsymbol{\Theta}_1 \boldsymbol{M}_1 \times_2 \cdots \times_K \boldsymbol{\Theta}_K \boldsymbol{M}_K, \tag{B.1}$$

where $\mathcal{Y} \in \mathbb{R}^{p_1 \times \cdots \times p_K}$ is the data tensor, $\mathcal{X} \in \mathbb{R}^{p_1 \times \cdots \times p_K}$ is the mean tensor, $\mathcal{S} \in \mathbb{R}^{r_1 \times \cdots \times r_K}$ is the core tensor, $\mathcal{E} \in \mathbb{R}^{p_1 \times \cdots \times p_K}$ is the noise tensor consisting of independent zero-mean sub-Gaussian entries with variance bounded by σ^2, $\boldsymbol{\Theta}_k = \mathrm{diag}(\boldsymbol{\theta}_k)$, and $\boldsymbol{M}_k \in \{0,1\}^{p_k \times r_k}$ is the membership matrix corresponding to the assignment $z_k : [p_k] \mapsto [r_k]$, for all $k \in [K]$.

For ease of notation, we use $\{z_k\}$ to denote the collection $\{z_k\}_{k=1}^K$, and $\{\boldsymbol{\theta}_k\}$ to denote the collection $\{\boldsymbol{\theta}_k\}_{k=1}^K$. Correspondingly, we consider the parameter space for the triplet $(\{z_k\}, \mathcal{S}, \{\boldsymbol{\theta}_k\})$,

$$\begin{aligned} \mathcal{P}(\{r_k\}) = \Big\{ & (\{z_k\}, \mathcal{S}, \{\boldsymbol{\theta}_k\}) : \boldsymbol{\theta}_k \in \mathbb{R}_+^p, \frac{c_1 p_k}{r_k} |z_k^{-1}(a)| \le \frac{c_2 p_k}{r_k}, \\ & c_3 \le \|\boldsymbol{S}_{k,a:}\| \le c_4, \|\boldsymbol{\theta}_{k, z_k^{-1}(a)}\|_1 = |z_k^{-1}(a)|, \\ & \text{for all } a \in [r_k], k \in [K] \Big\}. \end{aligned} \tag{B.2}$$

We call the degree heterogeneity $\{\boldsymbol{\theta}_k\}$ is balanced if for all $k \in [K]$,

$$\min_{a \in [r]} \|\boldsymbol{\theta}_{k, z_k^{-1}(a)}\| = (1 + o(1)) \max_{a \in [r]} \|\boldsymbol{\theta}_{k, z_k^{-1}(a)}\|.$$

We also consider the generalized Assumption 3.2 on angle gap.

Assumption B.1 (Generalized angle gap)**.** *Recall $\boldsymbol{S}_k = Mat_k(\mathcal{S})$. We assume the minimal gap between normalized rows of $\boldsymbol{S}_k$ is bounded away from zero for all $k \in [K]$; i.e.,*

$$\Delta_{\min} := \min_{k \in [K]} \min_{a \neq b \in [r_k]} \left\| \boldsymbol{S}_{k,a:}^s - \boldsymbol{S}_{k,b:}^s \right\| > 0.$$

Similarly, let SNR $= \Delta_{\min}^2/\sigma^2$ with the generalized minimal gap $\Delta_{\min}^2$ defined in Assumption B.1. We define the regime

$$\mathcal{P}(\gamma) = \mathcal{P}(\{r_k\}) \cap \{\mathcal{S} \text{ satisfies SNR} = p^\gamma \text{ and } p_k \asymp p, k \in [K]\}.$$

Proof of Theorem 3.3

Proof of Theorem 3.3. To study the identifiability, we consider the noiseless model with $\mathcal{E} = 0$. Assume that there exist two parameterizations satisfying

$$\begin{aligned} \mathcal{X} &= \mathcal{S} \times_1 \Theta_1 \boldsymbol{M}_1 \times_2 \cdots \times_K \Theta_K \boldsymbol{M}_K' \\ &= \mathcal{S}' \times_1 \Theta_1' \boldsymbol{M}_1' \times_2 \cdots \times_K \Theta_K' \boldsymbol{M}_K', \end{aligned} \tag{B.3}$$

where $(\{z_k\}, \mathcal{S}, \{\boldsymbol{\theta}_k\}) \in \mathcal{P}(\{r_k\})$ and $(\{z_k'\}, \mathcal{S}', \{\boldsymbol{\theta}_k'\}) \in \mathcal{P}(\{r_k'\})$ are two sets of parameters. We prove the sufficient and necessary conditions separately.

($\Leftarrow$) For the necessity, it suffices to construct two distinct parameters up to cluster label permutation, if the model (B.1) violates Assumption B.1. Note that $\Delta_{\min}^2 = 1$ when there exists $k \in [K]$ such that $r_k = 1$. Hence, we consider the case that $r_k \geq 2$ for all $k \in [K]$. Without loss of generality, we assume $\left\| \boldsymbol{S}_{1,1:}^s - \boldsymbol{S}_{1,2:}^s \right\| = 0$.

By constraints in parameter space (B.2), neither $\boldsymbol{S}_{1,1:}$ nor $\boldsymbol{S}_{1,2:}$ is a zero vector. There exists a positive constant c such that $\boldsymbol{S}_{1,1:} = c\boldsymbol{S}_{1,2:}$. Thus, there exists a core tensor $\mathcal{S}_0 \in \mathbb{R}^{r_1 - 1 \times \cdots \times r_K}$ such that

$$\mathcal{S} = \mathcal{S}_0 \times_1 \boldsymbol{C}\boldsymbol{R},$$

where $\boldsymbol{C} = \text{diag}(1, c, 1, ..., 1) \in \mathbb{R}^{r_1 \times r_1}$ and

$$\boldsymbol{R} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & \mathbf{1}_{r_1-2} \end{pmatrix} \in \mathbb{R}^{r_1 \times (r_1-1)}.$$

Let $\boldsymbol{D} = \text{diag}(1 + c, 1, ..., 1) \in \mathbb{R}^{r_1-1 \times r_1-1}$. Consider the parameterization $\boldsymbol{M}_1' = \boldsymbol{M}_1\boldsymbol{R}, \mathcal{S}' = \mathcal{S}_0 \times_1 \boldsymbol{D}$, and

$$\theta_1'(i) = \begin{cases} \frac{1}{1+c}\theta_1(i) & i \in z_1^{-1}(1), \\ \frac{c}{1+c}\theta_1(i) & i \in z_1^{-1}(2), \\ \theta_1(i) & \text{otherwise}, \end{cases}$$

and $\boldsymbol{M}_k' = \boldsymbol{M}_k, \boldsymbol{\theta}_k' = \boldsymbol{\theta}_k$ for all $k = 2, \ldots, K$. Then we have constructed a triplet $(\{z_k'\}, \mathcal{S}', \{\boldsymbol{\theta}_k'\})$ that is distinct from $(\{z_k\}, \mathcal{S}, \{\boldsymbol{\theta}_k\})$ up to label permutation.

($\Rightarrow$) For the sufficiency, it suffices to show that all possible triplets $(\{z_k'\}, \mathcal{S}', \{\boldsymbol{\theta}_k'\})$ are identical to $(\{z_k\}, \mathcal{S}, \{\boldsymbol{\theta}_k\})$ up to label permutation if the model (B.1) satisfies Assumption (B.1). We show the uniqueness of the three parameters, $\{\boldsymbol{M}_k\}, \{\mathcal{S}\}, \{\boldsymbol{\theta}_k\}$ separately.

First, we show the uniqueness of $\boldsymbol{M}_k$ for all $k \in [K]$. When $r_k = 1$, all possible $\boldsymbol{M}_k$'s are equal to the vector $\mathbf{1}_{p_k}$, and the uniqueness holds trivially. Hence, we consider the case that $r_k \geq 2$. Without loss of generality, we consider $k = 1$ with $r_1 \geq 2$ and show the uniqueness of the first mode membership matrix; i.e., $\boldsymbol{M}_1' = \boldsymbol{M}_1\boldsymbol{P}_1$ where $\boldsymbol{P}_1$ is a permutation matrix. The conclusion for $k \geq 2$ can be showed similarly and thus omitted.

Consider an arbitrary node pair (i, j). If $z_1(i) = z_1(j)$, then we have $\|\boldsymbol{X}^s_{1,z_1(i):} - \boldsymbol{X}^s_{1,z_1(j):}\| = 0$ and thus $\|(\boldsymbol{S}')^s_{1,z_1'(i):} - (\boldsymbol{S}')^s_{1,z_1'(j):}\| = 0$ by Lemma B.2. Then, by Assumption (B.1), we have $z_1'(i) = z_1'(j)$. Conversely, if $z_1(i) \neq z_1(j)$, then we have $\left\|\boldsymbol{X}^s_{1,i:} - \boldsymbol{X}^s_{1,j:}\right\| \neq 0$ and thus $\left\|(\boldsymbol{S}')^s_{1,z_1'(i):} - (\boldsymbol{S}')^s_{1,z_1'(j):}\right\| \neq 0$ by Lemma B.2.

Hence, we have $z_1'(i) \neq z_1'(j)$. Therefore, we have proven that z_1' is identical z_i up to label permutation.

Next, we show the uniqueness of $\boldsymbol{\theta}_k$ for all $k \in [K]$ provided that $z_k = z_k'$. Similarly, consider $k = 1$ only, and omit the procedure for $k \geq 2$.

Consider an arbitrary $j \in [p_1]$ such that $z_1(j) = a$. Then for all the nodes $i \in z_1^{-1}(a)$ in the same cluster of j, we have

$$\frac{\boldsymbol{X}_{1,z_1(i):}}{\boldsymbol{X}_{1,z_1(j):}} = \frac{\boldsymbol{X}'_{1,z_1(i):}}{\boldsymbol{X}'_{1,z_1(j):}}, \text{ which implies } \frac{\theta_1(j)}{\theta_1(i)} = \frac{\theta_1'(j)}{\theta_1'(i)}. \tag{B.4}$$

Let $\theta_1'(j) = c\theta_1(j)$ for some positive constant c. By equation (B.4), we have $\theta_1'(i) = c\theta_1(i)$ for all $i \in z_1^{-1}(a)$. By the constraint $(\{z_k\}, \mathcal{S}', \{\boldsymbol{\theta}_k'\}) \in \mathcal{P}(\{r_k\})$, we have

$$\sum_{j \in z_1^{-1}(a)} \theta_1'(j) = c \sum_{j \in z_1^{-1}(a)} \theta_1(j) = 1,$$

which implies $c = 1$. Hence, we have proven $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_1'$ provided that $z_1 = z_1'$.

Last, we show the uniqueness of $\mathcal{S}$; i.e., $\mathcal{S}' = \mathcal{S} \times_1 \boldsymbol{P}_1^{-1} \times_2 \cdots \times_K \boldsymbol{P}_K^{-1}$, where $\boldsymbol{P}_k$'s are permutation matrices for all $k \in [K]$. Provided $z_k' = z_k, \boldsymbol{\theta}_k' = \boldsymbol{\theta}_k$, we have $\boldsymbol{M}_k' = \boldsymbol{M}_k \boldsymbol{P}_k$ and $\boldsymbol{\Theta}_k' = \boldsymbol{\Theta}_k$ for all $k \in [K]$.

Let $\boldsymbol{D}_k = \left[(\boldsymbol{\Theta}_k' \boldsymbol{M}_k')^T (\boldsymbol{\Theta}_k' \boldsymbol{M}_k')\right]^{-1} (\boldsymbol{\Theta}_k' \boldsymbol{M}_k')^T, k \in [K]$. By the parameterization (B.3), we have

$$\begin{aligned}
\mathcal{S}' &= \mathcal{X} \times_1 \boldsymbol{D}_1 \times_2 \cdots \times_K \boldsymbol{D}_K \\
&= \mathcal{S} \times_1 \boldsymbol{D}_1 \boldsymbol{\Theta}_1 \boldsymbol{M}_1 \times_1 \cdots \times_K \boldsymbol{D}_K \boldsymbol{\Theta}_K \boldsymbol{M}_K \\
&= \mathcal{S} \times_1 \boldsymbol{P}_1^{-1} \times_2 \cdots \times_K \boldsymbol{P}_K^{-1}.
\end{aligned}$$

Therefore, we finish the proof of Theorem 3.3. □

Useful Lemma for the Proof of Theorem 3.3

Lemma B.2 (Motivation of angle-based clustering)**.** *Consider the signal tensor $\mathcal{X}$ in the general asymmetric dTBM* (B.1) *with $(\{z_k\}, \mathcal{S}, \{\boldsymbol{\theta}_k\}) \in \mathcal{P}(\{r_k\})$ and $r_k \geq 2, k \in [K]$.*

Then, for any $k \in [K]$ and index pair $(i, j) \in [p_k]^2$, we have

$$\left\|\boldsymbol{S}^s_{k,z_k(i):} - \boldsymbol{S}^s_{k,z_k(j):}\right\| = 0 \quad \textit{if and only if} \quad \left\|\boldsymbol{X}^s_{k,z_k(i):} - \boldsymbol{X}^s_{k,z_k(j):}\right\| = 0.$$

Proof of Lemma B.2. Without loss of generality, we prove $k = 1$ only and drop the subscript k in $\boldsymbol{X}_k, \boldsymbol{S}_k$ for notational convenience. By tensor matricization, we have

$$\boldsymbol{X}_{j:} = \theta_1(j)\boldsymbol{S}_{z_1(j):}\left[\boldsymbol{\Theta}_2\boldsymbol{M}_2 \otimes \cdots \otimes \boldsymbol{\Theta}_K\boldsymbol{M}_K\right]^T.$$

Let $\tilde{\boldsymbol{M}} = \boldsymbol{\Theta}_2\boldsymbol{M}_2 \otimes \cdots \otimes \boldsymbol{\Theta}_K\boldsymbol{M}_K$. Notice that for two vectors $\boldsymbol{a}, \boldsymbol{b}$ and two positive constants $c_1, c_2 > 0$, we have

$$\|\boldsymbol{a}^s - \boldsymbol{b}^s\| = \|(c_1\boldsymbol{a})^s - (c_2\boldsymbol{b})^s\|.$$

Thus it suffices to show the following statement holds for any index pair $(i, j) \in [p_1]^2$,

$$\left\|\boldsymbol{S}^s_{z_1(i):} - \boldsymbol{S}^s_{z_1(j):}\right\| = 0 \quad \text{if and only if} \quad \left\|\left[\boldsymbol{S}_{z_1(i):}\tilde{\boldsymbol{M}}^T\right]^s - \left[\boldsymbol{S}_{z_1(j):}\tilde{\boldsymbol{M}}^T\right]^s\right\| = 0.$$

($\Leftarrow$) Suppose $\left\|\left[\boldsymbol{S}_{z_1(i):}\tilde{\boldsymbol{M}}^T\right]^s - \left[\boldsymbol{S}_{z_1(j):}\tilde{\boldsymbol{M}}^T\right]^s\right\| = 0$. There exists a positive constant c such that $\boldsymbol{S}_{z_1(i):}\tilde{\boldsymbol{M}}^T = c\boldsymbol{S}_{z_1(j):}\tilde{\boldsymbol{M}}^T$. Note that

$$\boldsymbol{S}_{z_1(i):} = \boldsymbol{S}_{z_1(i):}\tilde{\boldsymbol{M}}^T\left[\tilde{\boldsymbol{M}}\left(\tilde{\boldsymbol{M}}^T\tilde{\boldsymbol{M}}\right)^{-1}\right],$$

where $\tilde{\boldsymbol{M}}^T\tilde{\boldsymbol{M}}$ is an invertiable diagonal matrix with positive diagonal elements. Thus, we have $\boldsymbol{S}_{z_1(i):} = c\boldsymbol{S}_{z_1(j):}$, which implies $\left\|\boldsymbol{S}^s_{z_1(i):} - \boldsymbol{S}^s_{z_1(j):}\right\| = 0$.

($\Rightarrow$) Suppose $\left\|\boldsymbol{S}^s_{z_1(i):} - \boldsymbol{S}^s_{z_1(j):}\right\| = 0$. There exists a positive constant c such that $\boldsymbol{S}_{z_1(i):} = c\boldsymbol{S}_{z_1(j):}$, and thus $\boldsymbol{S}_{z_1(i):}\tilde{\boldsymbol{M}}^T = c\boldsymbol{S}_{z_1(j):}\tilde{\boldsymbol{M}}^T$, which implies $\left\|\left[\boldsymbol{S}_{z_1(i):}\tilde{\boldsymbol{M}}^T\right]^s - \left[\boldsymbol{S}_{z_1(j):}\tilde{\boldsymbol{M}}^T\right]^s\right\| = 0$.

Therefore, we finish the proof of Lemma B.2. □

Proof of Lemma 3.4 and Lemma 3.11

Proof of Lemma 3.4. Note that the vector $\boldsymbol{S}_{z(i):}$ can be folded to a tensor $\mathcal{S}' = [\![\mathcal{S}'_{a_2,\ldots,a_K}]\!] \in \mathbb{R}^{r^{K-1}}$; i.e., $\mathrm{vec}(\mathcal{S}') = \boldsymbol{S}_{z(i):}$. Define weight vectors $\boldsymbol{w}_{a_2,\cdots,a_K}$ corresponding to the elements in $\mathcal{S}'_{a_2,\ldots,a_K}$ by

$$\boldsymbol{w}_{a_2\cdots a_K} = [\boldsymbol{\theta}^T_{z^{-1}(a_2)} \otimes \cdots \otimes \boldsymbol{\theta}^T_{z^{-1}(a_K)}] \in \mathbb{R}^{|z^{-1}(a_2)|\times\cdots\times|z^{-1}(a_K)|},$$

for all $a_k \in [r], k = 2, \ldots, K$, where $\otimes$ denotes the Kronecker product. Therefore, we have $\boldsymbol{X}_{i:} = \theta(i)\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{S}_{z(i):})$ where $\boldsymbol{w} = \{\boldsymbol{w}_{a_2,\cdots,a_K}\}_{a_k\in[r],k\in[K]/\{1\}}$. Specifically, we have $\|\boldsymbol{w}_{a_2,\ldots,a_K}\|^2 = \prod_{k=2}^K \|\boldsymbol{\theta}_{z^{-1}(a_k)}\|^2$, and by the balanced assumption (3.6) we have

$$\max_{(a_2,\ldots,a_K)} \|\boldsymbol{w}_{a_2,\ldots,a_K}\|^2 = (1+o(1)) \min_{(a_2,\ldots,a_K)} \|\boldsymbol{w}_{a_2,\ldots,a_K}\|^2. \tag{B.5}$$

Consider the inner product of $\boldsymbol{X}_{i:}$ and $\boldsymbol{X}_{j:}$ for $z(i) \neq z(j)$. By the definition of weighted padding operator (B.31) and the balanced assumption (B.5), we have

$$\begin{aligned}
&\langle \boldsymbol{X}_{i:}, \boldsymbol{X}_{j:} \rangle \\
&= \theta(i)\theta(j) \left\langle \mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{S}_{z(i):}), \mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{S}_{z(j):}) \right\rangle \\
&= \theta(i)\theta(j) \min_{(a_2,\ldots,a_K)} \|\boldsymbol{w}_{a_2,\ldots,a_K}\|^2 \left\langle \boldsymbol{S}_{z(i):}, \boldsymbol{S}_{z(j):} \right\rangle (1+o(1)).
\end{aligned}$$

Therefore, when p large enough, the inner product $\langle \boldsymbol{X}_{i:}, \boldsymbol{X}_{j:} \rangle$ has the same sign as $\left\langle \boldsymbol{S}_{z(i):}, \boldsymbol{S}_{z(j):} \right\rangle$.

Then, we have

$$\begin{aligned}
\cos(\boldsymbol{S}_{z_1(i):}, \boldsymbol{S}_{z_1(j):}) &= \frac{\left\langle \boldsymbol{S}_{z_1(i):}, \boldsymbol{S}_{z_1(j):} \right\rangle}{\|\boldsymbol{S}_{z_1(i):}\|\|\boldsymbol{S}_{z_1(j):}\|} \\
&= (1+o(1)) \frac{\langle \boldsymbol{X}_{i:}, \boldsymbol{X}_{j:} \rangle}{\|\boldsymbol{X}_{i:}\|\|\boldsymbol{X}_{j:}\|} \\
&= (1+o(1)) \cos(\boldsymbol{X}_{i:}, \boldsymbol{X}_{j:}),
\end{aligned}$$

where the second inequality follows by the balance assumption on $\boldsymbol{\theta}$.

Further, notice that $\|\boldsymbol{v}_1^s - \boldsymbol{v}_2^s\|^2 = 2(1-\cos(\boldsymbol{v}_1, \boldsymbol{v}_2))$. For all i, j such that $z(i) \neq z(j)$, when $p \to \infty$, we have

$$\|\boldsymbol{X}_{i:}^s - \boldsymbol{X}_{j:}^s\| \asymp \|\boldsymbol{S}_{z_1(i):}^s - \boldsymbol{S}_{z_1(j):}^s\| \gtrsim \Delta_{\min}.$$

□

Proof of Lemma 3.11. By the definition of minimal gap in Assumption 3.2, we have

$$\begin{aligned} L^{(t)} &= \frac{1}{p} \sum_{i\in[p]} \theta(i) \sum_{b\in[r]} \mathbb{1}\left\{z^{(t)}(i) = b\right\} \|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2 \\ &\geq \frac{1}{p} \sum_{i\in[p]} \theta(i) \sum_{b\in[r]} \mathbb{1}\left\{z^{(t)}(i) = b\right\} \Delta_{\min}^2 \\ &\geq c\ell^{(t)} \Delta_{\min}^2, \end{aligned}$$

where the last inequality follows from the assumption $\min_{i\in[p]} \theta(i) \geq c > 0$. □

Proof of Theorem 3.7 (Impossibility)

Proof of Theorem 3.7 (*Impossibility*). Consider the general asymmetric dTBM (B.1) in the special case that $p_k = p$ and $r_k = r$ for all $k \in [K]$ with $K \geq 2$, $2 \leq r \lesssim p^{1/3}$ as $p \to \infty$. For simplicity, we show the minimax rate for the estimation on the first mode $\hat{z}_1$; the proof for other modes are essentially the same.

To prove the minimax rate (3.10), it suffices to take an arbitrary $\mathcal{S}^* \in \mathcal{P}_{\mathcal{S}}(\gamma)$ wih $\gamma < -(K-1)$ and construct $(z_k^*, \boldsymbol{\theta}_k^*)$ such that

$$\inf_{\hat{z}_1} \mathbb{E}\left[p\ell(\hat{z}_1, z_1^*) | (z_k^*, \mathcal{S}^*, \boldsymbol{\theta}_k^*)\right] \geq 1.$$

We first define a subset of indices $T_k \subset [p_k], k \in [K]$ in order to avoid the complication of label permutation. Based on (Han et al., 2022a, Proof of Theorem 6), we consider the restricted family of $\hat{z}_k$'s for which the following three conditions

are satisfied:

$$\text{(a) } \hat{z}_k(i) = z_k(i) \text{ for all } i \in T_k; \quad \text{(b) } |T_k^c| \asymp \frac{p}{r};$$
$$\text{(c) } \min_{\pi \in \Pi} \sum_{i \in [p]} \mathbb{1}\{\hat{z}_k(i) \neq \pi \circ z_k(i)\} = \sum_{i \in [p]} \mathbb{1}\{\hat{z}_k(i) \neq z_k(i)\},$$

for all $k \in [K]$. Now, we consider the construction:

(i) $\{z_k^*\}$ satisfies properties (a)-(c) with misclassification sets T_k^c for all $k \in [K]$;

(ii) $\{\boldsymbol{\theta}_k^*\}$ such that $\boldsymbol{\theta}_k^*(i) \leq \sigma r^{(K-1)/2} p^{-(K-1)/2}$ for all $i \in T_k^c, k \in [K]$ and $\max_{k \in [K], a \in [r]} \|\boldsymbol{\theta}_{k, z_k^{*,-1}(a)}\|_2^2 \asymp p/r$.

Combining the inequalities (39) and (40) in the proof of Theorem 2 in Gao et al. (2018), we have

$$\begin{aligned} &\inf_{\hat{z}_1} \mathbb{E}\left[\ell(\hat{z}_1, z_1^*) | (z_k^*, \mathcal{S}^*, \boldsymbol{\theta}_k^*)\right] \geq \\ &\frac{C}{r^3 |T_1^c|} \sum_{i \in T_1^c} \inf_{\hat{z}_1(i)} \{\mathbb{P}[\hat{z}_1(i) = 1 | z_1^*(i) = 2, z_k^*, \mathcal{S}^*, \boldsymbol{\theta}_k^*] \\ &\qquad + \mathbb{P}[\hat{z}_1(i) = 2 | z_1^*(i) = 1, z_k^*, \mathcal{S}^*, \boldsymbol{\theta}_k^*]\}, \end{aligned} \tag{B.6}$$

where C is some positive constant, $\hat{z}_1$ on the left hand side denote the generic assignment functions in $\mathcal{P}(\gamma)$, and the infimum on the right hand side is taken over the generic assignment function family of $\hat{z}_1(i)$ for all nodes $i \in T_1^c$. Here, the factor $r^3 = r \cdot r^2$ in (B.6) comes from two sources: $r^2 \asymp \binom{r}{2}$ comes from the multiple testing burden for all pairwise comparisons among r clusters; and another r comes from the number of elements $|T_k^c| \asymp p/r$ to be clustered.

Next, we need to find the lower bound of the rightmost side in (B.6). We consider the hypothesis test based on model (B.1). First, we reparameterize the model under the construction (i)-(ii).

$$\boldsymbol{x}_a^* = \left[\text{Mat}_1\left(\mathcal{S}^* \times_2 \boldsymbol{\Theta}_2^* \boldsymbol{M}_2^* \times_3 \cdots \times_K \boldsymbol{\Theta}_K^* \boldsymbol{M}_K^*\right)\right]_{a:},$$

for all $a \in [r]$, where $\boldsymbol{x}_a^*$'s are centroids in $\mathbb{R}^{p^{K-1}}$. Without loss of generality, we consider the lower bound for the summand in (B.6) for $i = 1$. The analysis for other $i \in T_1^c$ are similar. For notational simplicity, we suppress the subscript i and write $\boldsymbol{y}, \theta^*, z$ in place of $\boldsymbol{y}_1, \boldsymbol{\theta}_1^*(1)$ and $z_1(1)$, respectively. The equivalent vector problem for assessing the summand in (B.6) is

$$\boldsymbol{y} = \theta^* \boldsymbol{x}_z^* + \boldsymbol{e}, \tag{B.7}$$

where $z \in \{1, 2\}$ is an unknown parameter, $\theta^* \in \mathbb{R}_+$ is the given heterogeneity degree, $\boldsymbol{x}_1^*, \boldsymbol{x}_2^* \in \mathbb{R}^{p^{K-1}}$ are given centroids, and $\boldsymbol{e} \in \mathbb{R}^{p^{K-1}}$ consists of i.i.d. $N(0, \sigma^2)$ entries. Then, we consider the hypothesis testing under the model (B.7):

$$H_0 : z = 1, \boldsymbol{y} = \theta^* \boldsymbol{x}_1^* + \boldsymbol{e} \;\leftrightarrow\; H_1 : z = 2, \boldsymbol{y} = \theta^* \boldsymbol{x}_2^* + \boldsymbol{e}, \tag{B.8}$$

The hypothesis testing (B.8) is a simple versus simple testing, since the assignment z is the only unknown parameter in the test. By Neyman-Pearson lemma, the likelihood ratio test is optimal with minimal Type I + II error. Under Gaussian model, the likelihood ratio test of (B.8) is equivalent to the least square estimator $\hat{z}_{LS} = \arg\min_{a=\{1,2\}} \|\boldsymbol{y} - \theta^* \boldsymbol{x}_a^*\|_F^2$.

Let $\boldsymbol{S} = \text{Mat}_1(\mathcal{S})$. Note that

$$\begin{aligned}
&\|\theta^* \boldsymbol{x}_1^* - \theta^* \boldsymbol{x}_2^*\|_F \\
&\leq \theta^* \|\boldsymbol{S}_{1:}^* - \boldsymbol{S}_{2:}^*\|_F \prod_{k=2}^{K} \lambda_{\max}(\boldsymbol{\Theta}_k^* \boldsymbol{M}_k^*) \\
&\leq \theta^* \|\boldsymbol{S}_{1:}^* - \boldsymbol{S}_{2:}^*\|_F \max_{k \in [K]/\{1\}, a \in [r]} \|\boldsymbol{\theta}_{k, z_k^{*,-1}(a)}\|_2^{K-1} \\
&\leq \sigma r^{(K-1)/2} p^{-(K-1)/2} 2 c_4 p^{(K-1)/2} r^{-(K-1)/2} \\
&\leq 2 c_4 \sigma,
\end{aligned}$$

where $\lambda_{\max}(\cdot)$ denotes the maximal singular value, the second inequality follows from Lemma B.6, and the third inequality follows from property (ii) and the boundedness constraint in $\mathcal{P}_{\mathcal{S}}(\gamma)$ such that $\|\boldsymbol{S}_{1:}^* - \boldsymbol{S}_{2:}^*\|_F \leq \|\boldsymbol{S}_{1:}^*\|_F + \|\boldsymbol{S}_{2:}^*\|_F \leq 2 c_4$.

Hence, we have

$$\begin{aligned}
&\inf_{\hat{z}_1(1)} \{\mathbb{P}[\hat{z}_1(1) = 1 | z_1^*(1) = 2, z_k^*, \mathcal{S}^*, \boldsymbol{\theta}_k^*] \\
&\qquad + \mathbb{P}[\hat{z}_1(1) = 2 | z_1^*(1) = 1, z_k^*, \mathcal{S}^*, \boldsymbol{\theta}_k^*]\} \\
&= 2\mathbb{P}[\hat{z}_{LS} = 1 | z_1^*(1) = 2, z_k^*, \mathcal{S}^*, \boldsymbol{\theta}_k^*] \\
&= 2\mathbb{P}[\|\boldsymbol{y} - \theta^* \boldsymbol{x}_1^*\|_F^2 \leq \|\boldsymbol{y} - \theta^* \boldsymbol{x}_2^*\|_F^2 | z_1^*(1) = 2, z_k^*, \mathcal{S}^*, \boldsymbol{\theta}_k^*] \\
&= 2\mathbb{P}[2\langle \boldsymbol{e}, \theta^* \boldsymbol{x}_1^* - \theta^* \boldsymbol{x}_2^* \rangle \geq \|\theta^* \boldsymbol{x}_1^* - \theta^* \boldsymbol{x}_2^*\|_F^2] \\
&= 2\mathbb{P}[N(0,1) \geq \theta^* \|\boldsymbol{x}_1^* - \boldsymbol{x}_2^*\|_F / (2\sigma)] \\
&\geq 2\mathbb{P}[N(0,1) \geq c_4] \geq c, \qquad \text{(B.9)}
\end{aligned}$$

where the first equation holds by symmetry, the third equation holds by rearrangement, the fourth equation holds from the fact that $\langle \boldsymbol{e}, \theta^* \boldsymbol{x}_1^* - \theta^* \boldsymbol{x}_2^* \rangle \sim N(0, \sigma \|\theta^* \boldsymbol{x}_1^* - \theta^* \boldsymbol{x}_2^*\|_F)$, and c is some positive constant in the last inequality.

Plugging the inequality (B.9) into the inequality (B.6) for all $i \in T_1^c$, then, we have

$$\liminf_{p \to \infty} \inf_{\hat{z}_1} \mathbb{E}\left[p\ell(\hat{z}_1, z_1^*) | z_k^*, \boldsymbol{\theta}_k^*, \mathcal{S}^*\right] \geq \liminf_{p \to \infty} \frac{Ccp}{r^3} \geq Cc,$$

where the last inequality follows by the condition $r = o(p^{1/3})$. By the discrete nature of the misclustering error, we obtain our conclusion

$$\liminf_{p \to \infty} \inf_{\mathcal{S}^* \in \mathcal{P}_{\mathcal{S}}(\gamma)} \inf_{\hat{z}_{\text{stat}}} \sup_{(z^*, \boldsymbol{\theta}^*) \in \mathcal{P}_{z,\boldsymbol{\theta}}} \mathbb{E}\left[p\ell(\hat{z}_{\text{stat}}, z)\right] \geq 1.$$

Last, with constructed $z_k^*, \boldsymbol{\theta}_k^*$ satisfying properties (i) and (ii) and $\gamma' < -(K-1)$, we construct a core tensor $\mathcal{S}^*$ such that $\Delta_{\boldsymbol{X}^*}^2 \leq p^{-(K-1)}$. Based on the property (ii) and the boundedness constraint of $\mathcal{S}^*$ in $\mathcal{P}$, we still have $\|\theta^* \boldsymbol{x}_1^* - \theta^* \boldsymbol{x}_2^*\|_F \leq 2c_4\sigma$. Hence, we obtain the desired result

$$\begin{aligned}
&\liminf_{p \to \infty} \inf_{\hat{z}_1} \sup_{(z, \mathcal{S}, \boldsymbol{\theta}) \in \mathcal{P}'(\gamma')} \mathbb{E}\left[p\ell(\hat{z}_1, z_1)\right] \\
&\quad \geq \liminf_{p \to \infty} \inf_{\hat{z}_{\text{stat}}} \mathbb{E}\left[p\ell(\hat{z}_1, z_1^*) | z_k^*, \mathcal{S}^*, \boldsymbol{\theta}_k^*\right] \geq 1.
\end{aligned}$$

□

Proof of Theorem 3.9 (Impossibility)

Proof of Theorem 3.9 (Impossibility). The idea of proving computational hardness is to show the computational lower bound for a special class of degree-corrected tensor clustering model with $K \geq 2$ and $r \geq 2$. We construct the following special class of higher-order degree-corrected tensor clustering model. For a given signal level $\gamma \in \mathbb{R}$ and noise variance σ, define a rank-2 symmetric tensor $\mathcal{S} \in \mathbb{R}^{3\times\cdots\times 3}$ subject to

$$\mathcal{S} = \mathcal{S}(\gamma) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}^{\otimes K} + \sigma p^{-\gamma/2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}^{\otimes K}. \tag{B.10}$$

Then, we consider the signal tensor family

$$\begin{aligned} \mathcal{P}_{\text{shifted}}(\gamma) = \{\mathcal{X} \colon \mathcal{X} = \mathcal{S} \times_1 \boldsymbol{M}_1 \times_2 \cdots \times_K \boldsymbol{M}_K, \text{ where} \\ \text{membership matrix } \boldsymbol{M}_k \in \{0,1\}^{p\times 3} \text{ satisfies} \\ |\boldsymbol{M}_k(:, i)| \asymp p \text{ for all } i \in [3] \text{ and } k \in [K]\}. \end{aligned}$$

We claim that the constructed family satisfies the following two properties:

(i) For every $\gamma \in \mathbb{R}$, $\mathcal{P}_{\text{shifted}}(\gamma) \subset \mathcal{P}(\gamma)$, where $\mathcal{P}(\gamma)$ is the degree-corrected cluster tensor family (3.5).

(ii) For every $\gamma \in \mathbb{R}$, $\{\mathcal{X} - 1 \colon \mathcal{X} \in \mathcal{P}_{\text{shifted}}(\gamma)\} \subset \mathcal{P}_{\text{non-degree}}(\gamma)$, where $\mathcal{P}_{\text{non-degree}}(\gamma)$ denotes the sub-family of rank-one tensor block model constructed in the proof of (Han et al., 2022a, Theorem 7).

The verification of the above two properties is provided in the end of this proof.

Now, following the proof of (Han et al., 2022a, Theorem 7), when $\gamma < -K/2$, every polynomial-time algorithm estimator $(\hat{\boldsymbol{M}}_k)_{k\in[K]}$ obeys

$$\liminf_{p\to\infty} \sup_{\mathcal{X}\in\mathcal{P}_{\text{non-degree}}(\gamma)} \mathbb{P}(\exists k \in [K],\ \hat{\boldsymbol{M}}_k \neq \boldsymbol{M}_k) \geq 1/2, \tag{B.11}$$

under the HPC Conjecture 3.4. The inequality (B.11) implies

$$\liminf_{p\to\infty} \sup_{\mathcal{X}\in\mathcal{P}_{\text{non-degree}}(\gamma)} \max_{k\in[K]} \mathbb{E}[p\ell(z_k, \hat{z}_k)] \geq 1.$$

Based on properties (i)-(ii), we conclude that

$$\liminf_{p\to\infty} \sup_{\mathcal{X}\in\mathcal{P}(\gamma)} \max_{k\in[K]} \mathbb{E}[p\ell(z_k, \hat{z}_k)] \geq 1.$$

We complete the proof by verifying the properties (i)-(ii). For (i), we verify that the angle gap for the core tensor $\mathcal{S}$ in (B.10) is on the order of $\sigma p^{-\gamma/2}$. Specifically, write $\mathbf{1} = (1,1,1)$ and $\boldsymbol{e} = (1,-1,0)$. We have

$$\text{Mat}(\mathcal{S}) = \begin{bmatrix} \text{Vec}(\mathbf{1}^{\otimes K-1}) + \sigma p^{-\gamma/2}\text{Vec}\left(\boldsymbol{e}^{\otimes(K-1)}\right) \\ \text{Vec}(\mathbf{1}^{\otimes K-1}) - \sigma p^{-\gamma/2}\text{Vec}\left(\boldsymbol{e}^{\otimes(K-1)}\right) \\ \text{Vec}(\mathbf{1}^{\otimes K-1}) \end{bmatrix}.$$

Based on the orthogonality $\langle \mathbf{1}, \boldsymbol{e}\rangle = 0$, the minimal angle gap among rows of $\text{Mat}(\mathcal{S})$ is

$$\begin{aligned} \Delta^2_{\min}(\mathcal{S}) &\asymp \tan^2(\text{Mat}(\mathcal{S})_{1:}, \text{Mat}(\mathcal{S})_{3:}) \\ &= \left(\frac{\|\boldsymbol{e}\|_2}{\|\mathbf{1}\|_2}\right)^{2(K-1)} \sigma^2 d^{-\gamma} \\ &\asymp \sigma^2 d^{-\gamma}. \end{aligned}$$

Therefore, we have shown that $\mathcal{P}_{\text{shifited}}(\gamma) = \mathcal{P}(\gamma)$. Finally, the property (ii) follows directly by comparing the definition of $\mathcal{S}$ in (B.10) with that in the proof of (Han

et al., 2022a, Theorem 7). □

Proof of Theorem 3.12 and Proposition 3.5

Proof of Theorem 3.12. We prove Theorem 3.12 under the dTBM (3.1) with symmetric mean tensor, parameters $(z, \mathcal{S}, \boldsymbol{\theta})$, fixed $r \geq 1, K \geq 2$, and i.i.d. noise. For the case $r = 1$, we have $L(z^{(0)}, z) = 0, \ell(z^{(0)}, z) = 0$ trivially. Hence, we focus on the proof of the first mode clustering $z_1^{(0)}$ with $r \geq 2$; the proofs for the other modes can be extended similarly. We drop the subscript k in the matricizations $\boldsymbol{M}_k, \boldsymbol{X}_k, \boldsymbol{S}_k$ and in the estimate $z_1^{(0)}$. We firstly show the proof with balanced $\boldsymbol{\theta}$.

We firstly show the upper bound for misclustering error $\ell(z^{(0)}, z)$.

First, by Lemma 3.4, there exists a positive constant such that $\min_{z(i)\neq z(j)} \left\|\boldsymbol{X}^s_{i:} - \boldsymbol{X}^s_{j:}\right\| \geq c_0\Delta_{\min}$. By the balance assumption on $\boldsymbol{\theta}$ and Lemma B.8, we have

$$\min_{\pi\in\Pi} \sum_{i:z^{(0)}(i)\neq\pi(z(i))} \theta(i)^2 \leq \sum_{i\in S_I} \theta(i)^2 + 4\sum_{i\in S} \theta(i)^2, \tag{B.12}$$

where

$$S_0 = \{i : \|\hat{\boldsymbol{X}}_{i:}\| = 0\}, S = \{i \in S_0^c : \|\hat{\boldsymbol{x}}_{z^{(0)}(i)} - \boldsymbol{X}^s_{i:}\| \geq c_0\Delta_{\min}/2\}.$$

On one hand, note that for any set $P \in [p]$,

$$\begin{aligned}
\sum_{i\in P}\|\boldsymbol{X}_{i:}\|^2 &= \sum_{i\in P}\|\theta(i)\boldsymbol{S}_{z(i):}(\boldsymbol{\Theta}\boldsymbol{M})^{T,\otimes(K-1)}\|^2 \\
&\geq \sum_{i\in P}\theta(i)^2 \min_{a\in[r]}\|\boldsymbol{S}_{a:}\|^2\lambda_r^{2(K-1)}(\boldsymbol{\Theta}\boldsymbol{M}) \\
&\gtrsim \sum_{i\in P}\theta(i)^2 p^{K-1}r^{-(K-1)},
\end{aligned}$$

where the last inequality follows Lemma B.6, the assumption that $\min_{i\in[p]}\theta(i) \geq c$, and the constraint $\min_{a\in[r]}\|\boldsymbol{S}_{a:}\| \geq c_3$ in the parameter space (3.2). Thus, we have

$$\sum_{i\in P}\theta(i)^2 \lesssim \sum_{i\in P}\|\boldsymbol{X}_{i:}\|^2 p^{-(K-1)}r^{K-1}. \tag{B.13}$$

On the other hand, note that

$$
\begin{aligned}
&\sum_{i\in S}\|\boldsymbol{X}_{i:}\|^2 \\
&\leq 2\sum_{i\in S}\|\hat{\boldsymbol{X}}_{i:}\|^2 + 2\sum_{i\in S}\|\hat{\boldsymbol{X}}_{i:}-\boldsymbol{X}_{i:}\|^2 && \text{(B.14)}\\
&\leq \frac{8}{c_0^2\Delta_{\min}^2}\sum_{i\in S}\|\hat{\boldsymbol{X}}_{i:}\|^2\|\hat{\boldsymbol{x}}_{z^{(0)}(i)}-\boldsymbol{X}^s_{i:}\|^2 + 2\|\hat{\mathcal{X}}-\mathcal{X}\|_F^2 && \text{(B.15)}\\
&\leq \frac{16}{c_0^2\Delta_{\min}^2}\sum_{i\in S}\|\hat{\boldsymbol{X}}_{i:}\|^2\left[\|\hat{\boldsymbol{x}}_{z^{(0)}(i)}-\hat{\boldsymbol{X}}^s_{i:}\|^2 + \|\hat{\boldsymbol{X}}^s_{i:}-\boldsymbol{X}^s_{i:}\|^2\right] \\
&\qquad\qquad + 2\|\hat{\mathcal{X}}-\mathcal{X}\|_F^2 && \text{(B.16)}\\
&\leq \frac{16(1+\eta)}{c_0^2\Delta_{\min}^2}\sum_{i\in S}\|\hat{\boldsymbol{X}}_{i:}\|^2\|\hat{\boldsymbol{X}}^s_{i:}-\boldsymbol{X}^s_{i:}\|^2 + 2\|\hat{\mathcal{X}}-\mathcal{X}\|_F^2 && \text{(B.17)}\\
&\leq \left(\frac{16(1+\eta)}{c_0^2\Delta_{\min}^2}+2\right)\|\hat{\mathcal{X}}-\mathcal{X}\|_F^2 && \text{(B.18)}\\
&\lesssim \left(\frac{16(1+\eta)}{c_0^2\Delta_{\min}^2}+2\right)\left(p^{K/2}r+pr^2+r^K\right)\sigma^2, && \text{(B.19)}
\end{aligned}
$$

where inequalities (B.14) and (B.16) follow from the triangle inequality, (B.15) follows from the definition of S, (B.17) follows from the update rule of k-means in Step 6 of Sub-algorithm 1, (B.18) follows from Lemma B.3, and the last inequality (B.19) follows from Lemma B.7. Also, note that

$$
\begin{aligned}
\sum_{i\in S_0}\|\boldsymbol{X}_{i:}\|^2 &= \sum_{i\in S_0}\|\hat{\boldsymbol{X}}_{i:}-\boldsymbol{X}_{i:}\|^2 \\
&\leq \|\hat{\mathcal{X}}-\mathcal{X}\|_F^2 \\
&\lesssim \left(p^{K/2}r+pr^2+r^K\right)\sigma^2, && \text{(B.20)}
\end{aligned}
$$

where the equation follows from the definition of S_0. Therefore, combining the

inequalities (B.12), (B.13), (B.19), and (B.20), we have

$$\begin{aligned}
&\min_{\pi\in\Pi} \sum_{i:z^{(0)}(i)\neq\pi(z(i))} \theta(i)^2 \\
&\lesssim \left(\sum_{i\in S}\|\boldsymbol{X}_{i:}\|^2 + \sum_{i\in S_0}\|\boldsymbol{X}_{i:}\|^2\right) p^{-(K-1)}r^{K-1} \\
&\lesssim \frac{\sigma^2 r^{K-1}}{\Delta_{\min}^2 p^{K-1}}\left(p^{K/2}r + pr^2 + r^K\right). \qquad \text{(B.21)}
\end{aligned}$$

With the assumption that $\min_{i\in[p]}\theta(i) \geq c$, we finally obtain the result

$$\ell(z^{(0)}, z) \lesssim \frac{1}{p}\min_{\pi\in\Pi}\sum_{i:z^{(0)}(i)\neq\pi(z(i))}\theta(i)^2 \lesssim \frac{r^K p^{-K/2}}{\text{SNR}},$$

where the last inequality follows from the definition $\text{SNR} = \Delta_{\min}^2/\sigma^2$.

Without the balanced $\boldsymbol{\theta}$, we have $\min_{z(i)\neq z(j)}\left\|\boldsymbol{X}_{i:}^s - \boldsymbol{X}_{j:}^s\right\| \geq c_0\Delta_{\boldsymbol{X}}$. Replacing the definition of S with $\Delta_{\boldsymbol{X}}$, we obtain the desired result.

Next, we show the bound for $L(z^{(0)}, z)$.

Note that $\boldsymbol{X}_{i:}^s$ have only r different values. We let $\boldsymbol{X}_a^s = \boldsymbol{X}_{i:}^s$ for all i such that $z(i) = a, a \in [r]$. Notice that

$$\|\boldsymbol{X}_{i:}\|^2 \gtrsim p^{K-1}r^{-(K-1)}$$

and

$$\|\boldsymbol{X}_{i:} - \hat{\boldsymbol{X}}_{i:}\|^2 \leq \|\hat{\mathcal{X}} - \mathcal{X}\|_F^2 \lesssim p^{K/2}r + pr^2 + r^K.$$

Therefore, when p is large enough, we have

$$\begin{aligned}
&\sum_{i\in[p]} \|\boldsymbol{X}_{i:}\|^2 \|\hat{\boldsymbol{X}}_i^s - \hat{\boldsymbol{x}}_{z^{(0)}(i)}\|^2 \\
&\lesssim \sum_{i\in[p]} \left(\|\boldsymbol{X}_{i:}\|^2 - \|\boldsymbol{X}_{i:} - \hat{\boldsymbol{X}}_{i:}\|^2\right) \|\hat{\boldsymbol{X}}_{i:}^s - \hat{\boldsymbol{x}}_{z^{(0)}(i)}\|^2 \\
&\lesssim \sum_{i\in[p]} \|\hat{\boldsymbol{X}}_{i:}\|^2 \|\hat{\boldsymbol{X}}_{i:}^s - \hat{\boldsymbol{x}}_{z^{(0)}(i)}\|^2 \\
&\lesssim \eta \sum_{i\in[p]} \|\hat{\boldsymbol{X}}_{i:}\|^2 \|\hat{\boldsymbol{X}}_{i:}^s - \boldsymbol{X}_{i:}^s\|^2 \\
&\lesssim \|\hat{\mathcal{X}} - \mathcal{X}\|_F^2 \\
&\lesssim p^{K/2} r + p r^2 + r^K. \qquad \text{(B.22)}
\end{aligned}$$

Hence, we have

$$\begin{aligned}
\sum_{i\in[p]} \|\hat{\boldsymbol{X}}_{i:}^s - \hat{\boldsymbol{x}}_{z^{(0)}(i)}\|^2 &\lesssim \sum_{i\in[p]} \theta(i)^2 \|\hat{\boldsymbol{X}}_i^s - \hat{\boldsymbol{x}}_{z^{(0)}(i)}\|^2 \\
&\lesssim \frac{r^{K-1}}{p^{K-1}} \sum_{i\in[p]} \|\boldsymbol{X}_{i:}\|^2 \|\hat{\boldsymbol{X}}_{i:}^s - \hat{\boldsymbol{x}}_{z^{(0)}(i)}\|^2 \\
&\lesssim \frac{r^{K-1}}{p^{K-1}} \left(p^{K/2} r + p r^2 + r^K\right), \qquad \text{(B.23)}
\end{aligned}$$

where the first inequality follows from the assumption $\min_{i\in[p]} \theta(i) \geq c > 0$, the second inequality follows from the inequality (B.13), and the last inequality comes from the inequality (B.22).

Next, we consider the following quantity,

$$\begin{aligned}
&\sum_{i\in[p]} \theta(i)\|\boldsymbol{X}^s_{i:} - \hat{\boldsymbol{x}}_{z^{(0)}(i)}\|^2 \\
&\lesssim \sum_{i\in[p]} \theta(i)^2\|\boldsymbol{X}^s_{i:} - \hat{\boldsymbol{X}}^s_{i:}\|^2 + \sum_{i\in[p]} \theta(i)^2\|\hat{\boldsymbol{X}}^s_{i:} - \hat{\boldsymbol{x}}_{z^{(0)}(i)}\|^2 \\
&\lesssim \sum_{i\in[p]} \frac{\theta(i)^2}{\|\boldsymbol{X}_{i:}\|^2}\|\boldsymbol{X}_{i:} - \hat{\boldsymbol{X}}_{i:}\|^2 + \sum_{i\in[p]} \theta(i)^2\|\hat{\boldsymbol{X}}^s_{i:} - \hat{\boldsymbol{x}}_{z^{(0)}(i)}\|^2 \\
&\lesssim \frac{r^{K-1}}{p^{K-1}}\left(p^{K/2}r + pr^2 + r^K\right), \qquad \text{(B.24)}
\end{aligned}$$

where the first inequality follows from the assumption of $\theta(i)$ and triangle inequality, the second inequality follows from Lemma B.3, and the last inequality follows from (B.23). In addition, with Theorem 3.12 and the condition SNR $\gtrsim p^{-K/2}\log p$, for all $a \in [r]$, we have

$$|z^{-1}(a) \cap (z^{(0)})^{-1}(a)| \geq |z^{-1}(a)| - p\ell(z^{(0)}, z) \gtrsim \frac{p}{r} - \frac{p}{\log p} \gtrsim \frac{p}{r},$$

when p is large enough. Therefore, for all $a \in [r]$, we have

$$\begin{aligned}
\|\hat{\boldsymbol{x}}_a - \boldsymbol{X}^s_a\|^2 &= \frac{\sum_{i\in z^{-1}(a)\cap(z^{(0)})^{-1}(a)} \left\|\boldsymbol{X}^s_{i:} - \hat{\boldsymbol{x}}_{z^{(0)}(i)}\right\|^2}{|z^{-1}(a) \cap (z^{(0)})^{-1}(a)|} \\
&\lesssim \frac{r}{p}\left(\sum_{i\in[p]}\|\boldsymbol{X}^s_{i:} - \hat{\boldsymbol{X}}^s_{i:}\|^2 + \sum_{i\in[p]}\|\hat{\boldsymbol{X}}^s_{i:} - \hat{\boldsymbol{x}}_{z^{(0)}(i)}\|^2\right) \\
&\lesssim \frac{r^K}{p^K}\left(p^{K/2}r + pr^2 + r^K\right), \qquad \text{(B.25)}
\end{aligned}$$

where the last inequality follows from the inequality (B.23).

Finally, we obtain

$$\begin{aligned}
L^{(0)} &= \frac{1}{p}\sum_{i\in[p]}\theta(i)\sum_{b\in[r]}\mathbb{1}\left\{z^{(0)}(i)=b\right\}\|[\boldsymbol{S}_{z(i):}]^s-[\boldsymbol{S}_{b:}]^s\|^2\\
&\lesssim \frac{1}{p}\sum_{i\in[p],z^{(0)}(i)\neq z(i)}\theta(i)\|\boldsymbol{X}^s_{i:}-\boldsymbol{X}^s_{z^{(0)}(i)}\|^2\\
&\lesssim \frac{1}{p}\sum_{i\in[p],z^{(0)}(i)\neq z(i)}\theta(i)\Big(\|\boldsymbol{X}^s_{i:}-\hat{\boldsymbol{x}}_{z^{(0)}(i)}\|^2\\
&\qquad\qquad + \|\hat{\boldsymbol{x}}_{z^{(0)}(i)}-\boldsymbol{X}^s_{z^{(0)}(i)}\|^2\Big)\\
&\leq \bar{C}\frac{r^K}{p^K}\left(p^{K/2}r+pr^2+r^K\right),\\
&\leq \frac{\bar{C}\Delta^2_{\min}}{\tilde{C}r\log p}
\end{aligned}$$

where the first inequality follows from Lemma 3.4, the third inequality follows from inequalities (B.24) and (B.25), and the last inequality follows from the assumption that SNR $\geq \tilde{C}p^{-K/2}\log p$. □

Proof of Proposition 3.5. Sub-algorithm 3 shares the same algorithm strategy as Sub-algorithm 1 but with a different estimation of the mean tensor, $\hat{\mathcal{X}}'$. Hence, the proof of Proposition 3.5 follows the same proof idea with the proof of Theorem 3.12. Replacing the estimation $\hat{\mathcal{X}}$ by $\hat{\mathcal{X}}'$ in the proof of Theorem 3.12, we have

$$\begin{aligned}
&\min_{\pi\in\Pi}\sum_{i:z^{(0)}(i)\neq\pi(z(i))}\theta(i)^2\\
&\quad\lesssim \left(\sum_{i\in S}\|\boldsymbol{X}_{i:}\|^2+\sum_{i\in S_0}\|\boldsymbol{X}_{i:}\|^2\right)p^{-(K-1)}r^{K-1}. \qquad \text{(B.26)}
\end{aligned}$$

By inequalities (B.18) and (B.20), we have

$$\sum_{i\in S}\|\boldsymbol{X}_{i:}\|^2 \le \left(\frac{16(1+\eta)}{c_0^2\Delta_{\min}^2}+2\right)\|\hat{\mathcal{X}}'-\mathcal{X}\|_F^2, \tag{B.27}$$

$$\sum_{i\in S_0}\|\boldsymbol{X}_{i:}\|^2 \le \|\hat{\mathcal{X}}'-\mathcal{X}\|_F^2. \tag{B.28}$$

Hence, it suffices to find the upper bound of the estimation error $\|\hat{\mathcal{X}}'-\mathcal{X}\|_F^2$ to complete our proof. Note that the matricization $\text{Mat}_{sq}(\mathcal{X}) \in \mathbb{R}^{p^{\lfloor K/2\rfloor}\times p^{\lceil K/2\rceil}}$ has $\text{rank}(\text{Mat}_{sq}(\mathcal{X})) \le r^{\lceil K/2\rceil}$, and Bernoulli random variables follow the sub-Gaussian distribution with bounded variance $\sigma^2 = 1/4$. Apply Lemma B.9 to $\boldsymbol{Y} = \text{Mat}_{sq}(\mathcal{Y})$, $\boldsymbol{X} = \text{Mat}_{sq}(\mathcal{X})$, and $\hat{\boldsymbol{X}} = \text{Mat}_{sq}(\hat{\mathcal{X}}')$. Then, with probability tending to 1 as $p\to\infty$, we have

$$\|\hat{\mathcal{X}}'-\mathcal{X}\|_F^2 = \|\text{Mat}_{sq}(\hat{\mathcal{X}}')-\text{Mat}_{sq}(\mathcal{X})\|_F^2 \lesssim p^{\lceil K/2\rceil}. \tag{B.29}$$

Combining the estimation error (B.29) with inequalities (B.27), (B.28), and (B.26), we obtain

$$\min_{\pi\in\Pi}\sum_{i:z^{(0)}(i)\neq\pi(z(i))}\theta(i)^2 \lesssim \frac{\sigma^2 r^{K-1}}{\Delta_{\min}^2 p^{K-1}}p^{\lceil K/2\rceil}. \tag{B.30}$$

Replace the inequality (B.21) in the proof of Theorem 3.12 by inequality (B.30). With the the same procedures to obtain $\ell(\hat{z}^{(0)},z)$ and $L(\hat{z}^{(0)},z)$ for Theorem 3.12 , we finish the proof of Proposition 3.5. □

Useful Definitions and Lemmas for the Proof of Theorem 3.12

Lemma B.3 (Basic inequality)**.** *For any two nonzero vectors $\boldsymbol{v}_1,\boldsymbol{v}_2$ of same dimension, we have*

$$\sin(\boldsymbol{v}_1,\boldsymbol{v}_2) \le \|\boldsymbol{v}_1^s-\boldsymbol{v}_2^s\| \le \frac{2\|\boldsymbol{v}_1-\boldsymbol{v}_2\|}{\max(\|\boldsymbol{v}_1\|,\|\boldsymbol{v}_2\|)}.$$

Proof of Lemma B.3. For the first inequality, let $\alpha\in[0,\pi]$ denote the angle between $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$. We have

$$\|\boldsymbol{v}_1^s-\boldsymbol{v}_2^s\| = \sqrt{2(1-\cos\alpha)} = 2\sin\frac{\alpha}{2} \ge \sin\alpha,$$

where the equations follow from the properties of trigonometric function and the inequality follows from the fact the $\cos\frac{\alpha}{2} \leq 1$ and $\sin\alpha = 2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2} > 0$ for $\alpha \in [0, \pi]$.

For the second inequality, without loss of generality, we assume $\|\boldsymbol{v}_1\| \geq \|\boldsymbol{v}_2\|$. Then

$$\begin{aligned}\|\boldsymbol{v}_1^s - \boldsymbol{v}_2^s\| &= \left\|\frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|} - \frac{\boldsymbol{v}_2}{\|\boldsymbol{v}_1\|} + \frac{\boldsymbol{v}_2}{\|\boldsymbol{v}_1\|} - \frac{\boldsymbol{v}_2}{\|\boldsymbol{v}_2\|}\right\| \\ &\leq \frac{\|\boldsymbol{v}_1 - \boldsymbol{v}_2\|}{\|\boldsymbol{v}_1\|} + \frac{\|\boldsymbol{v}_2\|\,\|\boldsymbol{v}_1\| - \|\boldsymbol{v}_2\|}{\|\boldsymbol{v}_1\|\,\|\boldsymbol{v}_2\|} \\ &\leq \frac{2\,\|\boldsymbol{v}_1 - \boldsymbol{v}_2\|}{\|\boldsymbol{v}_2\|}.\end{aligned}$$

Therefore, Lemma B.3 is proved. □

Definition B.4 (Weighted padding vectors)**.** *For a vector $\boldsymbol{a} = [\![a_i]\!] \in \mathbb{R}^d$, we define the padding vector of $\boldsymbol{a}$ with the weight collection $\boldsymbol{w} = \{\boldsymbol{w}_i \colon \boldsymbol{w}_i = [\![w_{ik}]\!] \in \mathbb{R}^{p_i}\}_{i=1}^d$ as*

$$\mathit{Pad}_{\boldsymbol{w}}(\boldsymbol{a}) = [a_1 \circ \boldsymbol{w}_1, \ldots, a_d \circ \boldsymbol{w}_d]^T, \tag{B.31}$$

where $a_i \circ \boldsymbol{w}_i = [a_i w_{i1}, \ldots, a_i w_{ip_i}]^T$, for all $i \in [d]$. Here we also view $\mathit{Pad}_{\boldsymbol{w}}(\cdot) \colon \mathbb{R}^d \mapsto \mathbb{R}^{\sum_{i\in[d]} p_i}$ as an operator. We have the bounds of the weighted padding vector

$$\min_{i\in[d]} \|\boldsymbol{w}_i\|^2 \|\boldsymbol{a}\|^2 \leq \|\mathit{Pad}_{\boldsymbol{w}}(\boldsymbol{a})\|^2 \leq \max_{i\in[d]} \|\boldsymbol{w}_i\|^2 \|\boldsymbol{a}\|^2. \tag{B.32}$$

Further, we define the inverse weighted padding operator $\mathit{Pad}^{-1} : \mathbb{R}^{\sum_{i\in[d]} p_i} \mapsto \mathbb{R}^d$ which satisfies

$$\mathit{Pad}_{\boldsymbol{w}}^{-1}(\mathit{Pad}_{\boldsymbol{w}}(\boldsymbol{a})) = \boldsymbol{a}.$$

Lemma B.5 (Angle for weighted padding vectors)**.** *Suppose that we have two non-zero*

vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$. Given the weight collection $\boldsymbol{w}$, we have

$$\frac{\min_{i\in[d]}\|\boldsymbol{w}_i\|}{\max_{i\in[d]}\|\boldsymbol{w}_i\|}\sin(\boldsymbol{a},\boldsymbol{b}) \overset{*}{\leq} \sin(\mathit{Pad}_{\boldsymbol{w}}(\boldsymbol{a}), \mathit{Pad}_{\boldsymbol{w}}(\boldsymbol{b})) \overset{**}{\leq} \frac{\max_{i\in[d]}\|\boldsymbol{w}_i\|}{\min_{i\in[d]}\|\boldsymbol{w}_i\|}\sin(\boldsymbol{a},\boldsymbol{b}). \tag{B.33}$$

Proof of Lemma B.5. We prove the two inequalities separately with similar ideas.

First, we prove the inequality ** in (B.33). Decomposing $\boldsymbol{b}$ yields

$$\boldsymbol{b} = \cos(\boldsymbol{a},\boldsymbol{b})\frac{\|\boldsymbol{b}\|}{\|\boldsymbol{a}\|}\boldsymbol{a} + \sin(\boldsymbol{a},\boldsymbol{b})\frac{\|\boldsymbol{b}\|}{\|\boldsymbol{a}^\perp\|}\boldsymbol{a}^\perp,$$

where $\boldsymbol{a}^\perp \in \mathbb{R}^d$ is in the orthogonal complement space of $\boldsymbol{a}$. By the Definition B.4, we have

$$\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{b}) = \cos(\boldsymbol{a},\boldsymbol{b})\frac{\|b\|}{\|a\|}\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{a}) + \sin(\boldsymbol{a},\boldsymbol{b})\frac{\|b\|}{\|a^\perp\|}\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{a}^\perp).$$

Note that $\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{a}^\perp)$ is not necessary equal to the orthogonal vector of $\mathrm{Pad}(\boldsymbol{a})$; i.e., $\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{a}^\perp) \neq (\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{a}))^\perp$. By the geometry property of trigonometric functions, we obtain

$$\begin{aligned}\sin(\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{a}), \mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{b})) &\leq \frac{\|\boldsymbol{b}\|\|\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{a}^\perp)\|}{\|\boldsymbol{a}^\perp\|\|\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{b})\|}\sin(\boldsymbol{a},\boldsymbol{b}) \\ &\leq \frac{\max_{i\in[d]}\|\boldsymbol{w}_i\|}{\min_{i\in[d]}\|\boldsymbol{w}_i\|}\sin(\boldsymbol{a},\boldsymbol{b}),\end{aligned}$$

where the second inequality follows by applying the property (B.32) to vectors $\boldsymbol{b}$ and $\boldsymbol{a}^\perp$.

Next, we prove inequality * in (B.33). With the decomposition of $\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{b})$ and the inverse weighted padding operator, we have

$$\boldsymbol{b} = \cos(\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{a}), \mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{b}))\frac{\|\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{b})\|}{\|\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{a})\|}\boldsymbol{a} + \sin(\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{a}), \mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{b}))\frac{\|\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{b})\|}{\|(\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{a}))^{\perp}\|}\mathrm{Pad}_{\boldsymbol{w}}^{-1}((\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{a}))^{\perp}).$$

Therefore, we obtain

$$\begin{aligned}
&\sin(\boldsymbol{a}, \boldsymbol{b}) \\
&\leq \frac{\|\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{b})\|\|\mathrm{Pad}_{\boldsymbol{w}}^{-1}((\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{a}))^{\perp})\|}{\|(\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{a}))^{\perp}\|\|\boldsymbol{b}\|}\sin(\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{a}), \mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{b})) \\
&\leq \frac{\max_{i\in[d]}\|\boldsymbol{w}_i\|}{\min_{i\in[d]}\|\boldsymbol{w}_i\|}\sin(\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{a}), \mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{b})),
\end{aligned}$$

where the second inequality follows by applying the property (B.32) to vectors $\boldsymbol{b}$ and $\mathrm{Pad}_{\boldsymbol{w}}^{-1}((\mathrm{Pad}_{\boldsymbol{w}}(\boldsymbol{a}))^{\perp})$. □

Lemma B.6 (Singular value of weighted membership matrix)**.** *Under the parameter space* (3.2) *and assumption that* $\min_{i\in[p]}\theta(i) \geq c$ *for some constant* $c > 0$*, the singular values of* $\boldsymbol{\Theta M}$ *are bounded as*

$$\begin{aligned}
\sqrt{p/r} &\lesssim \sqrt{\min_{a\in[r]}\|\boldsymbol{\theta}_{z^{-1}(a)}\|^2} \leq \lambda_r(\boldsymbol{\Theta M}) \\
&\leq \|\boldsymbol{\Theta M}\|_{\sigma} \leq \sqrt{\max_{a\in[r]}\|\boldsymbol{\theta}_{z^{-1}(a)}\|^2} \lesssim p/r.
\end{aligned}$$

Proof of Lemma B.6. Note that

$$(\boldsymbol{\Theta M})^T\boldsymbol{\Theta M} = \boldsymbol{D},$$

with $\boldsymbol{D} = \mathrm{diag}(D_1, \ldots, D_r)$ where $D_a = \|\boldsymbol{\theta}_{z^{-1}(a)}\|^2, a \in [r]$. By the definition of singular values, we have

$$\sqrt{\min_{a\in[r]}\|\boldsymbol{\theta}_{z^{-1}(a)}\|^2} \leq \lambda_r(\boldsymbol{\Theta M}) \leq \|\boldsymbol{\Theta M}\|_{\sigma} \leq \sqrt{\max_{a\in[r]}\|\boldsymbol{\theta}_{z^{-1}(a)}\|^2}.$$

Since that $\min_{i\in[p]}\theta(i)\geq c$ by the constraints in parameter space, we have

$$\min_{a\in[r]}\|\boldsymbol{\theta}_{z^{-1}(a)}\|^2\geq c^2\min_{a\in[r]}|z^{-1}(a)|\gtrsim\frac{p}{r},$$

where the last inequality follows from the constraint in parameter space (3.2). Finally, notice that

$$\sqrt{\max_{a\in[r]}\|\boldsymbol{\theta}_{z^{-1}(a)}\|^2}\leq\max_{a\in[r]}\sqrt{\|\boldsymbol{\theta}_{z^{-1}(a)}\|_1^2}\lesssim\frac{p}{r}.$$

Therefore, we complete the proof of Lemma B.6. □

Lemma B.7 (Singular-value gap-free tensor estimation error bound)**.** *Consider an order-K tensor $\mathcal{A}=\mathcal{X}+\mathcal{Z}\in\mathbb{R}^{p\times\cdots\times p}$, where $\mathcal{X}$ has Tucker rank $(r,...r)$ and $\mathcal{Z}$ has independent sub-Gaussian entries with parameter σ^2. Let $\hat{\mathcal{X}}$ denote the double projection estimated tensor in Step 2 of Sub-algorithm 1 in the main paper. Then with probability at least $1-C\exp(-cp)$, we have*

$$\|\hat{\mathcal{X}}-\mathcal{X}\|_F^2\leq C\sigma^2\left(p^{K/2}r+pr^2+r^K\right),$$

where C,c are some positive constants.

Proof of Lemma B.7. See (Han et al., 2022a, Proposition 1). □

Lemma B.8 (Upper bound of misclustering error)**.** *Let $z:[p]\mapsto[r]$ be a cluster assignment such that $|z^{-1}(a)|\asymp p/r$ for all $a\in[r]$ with $r\geq 2$. Let node i correspond to a vector $\boldsymbol{x}_i=\theta(i)\boldsymbol{v}_{z(i)}\in\mathbb{R}^d$, where $\{\boldsymbol{v}_a\}_{a=1}^r$ are the cluster centers and $\boldsymbol{\theta}=[\![\theta(i)]\!]\in\mathbb{R}_+^p$ is the positive degree heterogeneity. Assume that $\boldsymbol{\theta}$ satisfies the balanced assumption (3.6) such that $\frac{\max_{a\in[r]}\|\boldsymbol{\theta}_{z^{-1}(a)}\|^2}{\min_{a\in[r]}\|\boldsymbol{\theta}_{z^{-1}(a)}\|^2}=1+o(1)$. Consider an arbitrary estimate $\hat{z}$ with $\hat{\boldsymbol{x}}_i=\hat{\boldsymbol{v}}_{\hat{z}(i)}$ for all $i\in S$. Then, if*

$$\min_{a\neq b\in[r]}\|\boldsymbol{v}_a-\boldsymbol{v}_b\|\geq 2c, \tag{B.34}$$

for some constant $c > 0$, we have

$$\min_{\pi\in\Pi} \sum_{i:\hat{z}(i)\neq\pi(z(i))} \theta(i)^2 \leq \sum_{i\in S_0} \theta(i)^2 + 4\sum_{i\in S} \theta(i)^2,$$

where S_0 is defined in Step 4 of Sub-algorithm 1 and

$$S = \{i \in S_0^c : \|\hat{\boldsymbol{x}}_i - \boldsymbol{v}_{z(i)}\| \geq c\}.$$

Proof of Lemma B.8. For each cluster $u \in [r]$, we use C_u to collect the subset of points for which the estimated and true positions $\hat{\boldsymbol{x}}_i, \boldsymbol{x}_i$ are within distance c. Specifically, define

$$C_u = \{i \in z^{-1}(u) \cap S_0^c : \|\hat{\boldsymbol{x}}_i - \boldsymbol{v}_{z(i)}\| < c\},$$

and divide $[r]$ into three groups based on C_u as

$$\begin{aligned}
R_1 &= \{u \in [r] : C_u = \emptyset\},\\
R_2 &= \{u \in [r] : C_u \neq \emptyset, \text{ for all } i, j \in C_u, \hat{z}(i) = \hat{z}(j)\},\\
R_3 &= \{u \in [r] : C_u \neq \emptyset, \text{ there exist } i, j \in C_u, \hat{z}(i) \neq \hat{z}(j)\}.
\end{aligned}$$

Note that $\cup_{u\in[r]} C_u = S_0^c/S^c$ and $C_u \cap C_v = \emptyset$ for any $u \neq v$. Suppose there exist $i \in C_u$ and $j \in C_v$ with $u \neq v \in [r]$ and $\hat{z}(i) = \hat{z}(j)$. Then we have

$$\|\boldsymbol{v}_{z(i)} - \boldsymbol{v}_{z(j)}\| \leq \|\boldsymbol{v}_{z(i)} - \hat{\boldsymbol{x}}_i\| + \|\boldsymbol{v}_{z(j)} - \hat{\boldsymbol{x}}_j\| < 2c,$$

which contradicts to the assumption (B.34). Hence, the estimates $\hat{z}(i) \neq \hat{z}(j)$ for the nodes $i \in C_u$ and $j \in C_v$ with $u \neq v$. By the definition of R_2, the nodes in $\cup_{u\in R_2} C_u$ have the same assignment with z and $\hat{z}$. Then, we have

$$\min_{\pi\in\Pi} \sum_{i:\hat{z}(i)\neq\pi(z(i))} \theta(i)^2 \leq \sum_{i\in S_0} \theta(i)^2 + \sum_{i\in S} \theta(i)^2 + \sum_{i\in\cup_{u\in R_3} C_u} \theta(i)^2.$$

We only need to bound $\sum_{i\in\cup_{u\in R_3} C_u} \theta(i)^2$ to finish the proof. Note that every C_u with $u \in R_3$ contains at least two nodes assigned to different clusters by $\hat{z}$. Then, we have $|R_2| + 2|R_3| \leq r$. Since $|R_1| + |R_2| + |R_3| = r$, we have $|R_3| \leq |R_1|$. Hence, we

obtain

$$\begin{aligned}\sum_{i\in\cup_{u\in R_3}C_u}\theta(i)^2 &\le |R_3|\max_{a\in[r]}\|\boldsymbol{\theta}_{z^{-1}(a)}\|^2\\ &\le |R_1|\max_{a\in[r]}\|\boldsymbol{\theta}_{z^{-1}(a)}\|^2\\ &\le \frac{\max_{a\in[r]}\|\boldsymbol{\theta}_{z^{-1}(a)}\|^2}{\min_{a\in[r]}\|\boldsymbol{\theta}_{z^{-1}(a)}\|^2}\sum_{i\in\cup_{u\in R_1}z^{-1}(u)}\theta(i)^2\\ &\le 2\sum_{i\in S}\theta(i)^2,\end{aligned}$$

where the last inequality holds by the balanced assumption on $\boldsymbol{\theta}$ when p is large enough, and the fact that $\cup_{u\in R_1}z^{-1}(u)\subset S$. □

Lemma B.9 (Low-rank matrix estimation)**.** *Let $\boldsymbol{Y}=\boldsymbol{X}+\boldsymbol{E}\in\mathbb{R}^{m\times n}$, where $n>m$ and $\boldsymbol{E}$ contains independent mean-zero sub-Gaussian entries with bounded variance σ^2. Suppose rank$(\boldsymbol{X})=r$. Consider the least square estimator*

$$\hat{\boldsymbol{X}}=\underset{\boldsymbol{X}'\in\mathbb{R}^{m\times n},\mathit{rank}(\boldsymbol{X}')\le r}{\arg\min}\|\boldsymbol{X}'-\boldsymbol{Y}\|_F^2.$$

There exist positive constants C_1, C_2 such that

$$\|\hat{\boldsymbol{X}}-\boldsymbol{X}\|_F^2\le C_1\sigma^2 nr,$$

with probability at least $1-\exp(-C_2 nr)$.

Proof of Lemma B.9. Note that $\|\hat{\boldsymbol{X}}-\boldsymbol{Y}\|_F^2\le\|\boldsymbol{X}-\boldsymbol{Y}\|_F^2$ by the definition of least square estimator.

We have

$$\begin{aligned}&\|\hat{\boldsymbol{X}}-\boldsymbol{X}\|_F^2\\ &\le 2\left\langle\hat{\boldsymbol{X}}-\boldsymbol{X},\boldsymbol{Y}-\boldsymbol{X}\right\rangle\\ &\le 2\|\hat{\boldsymbol{X}}-\boldsymbol{X}\|_F\sup_{\boldsymbol{T}\in\mathbb{R}^{m\times n},\mathrm{rank}(\boldsymbol{T})\le 2r,\|\boldsymbol{T}\|_F=1}\langle\boldsymbol{T},\boldsymbol{Y}-\boldsymbol{X}\rangle\end{aligned}\tag{B.35}$$

with probability at least $1 - \exp(-C_2 nr)$, where the second inequality follows by re-arrangement.

Consider the SVD for matrix $\boldsymbol{T} = \boldsymbol{U}\Sigma\boldsymbol{V}^T$ with orthogonal matrices $\boldsymbol{U} \in \mathbb{R}^{m\times 2r}, \boldsymbol{V} \in \mathbb{R}^{n\times 2r}$ and diagonal matrix $\Sigma \in \mathbb{R}^{2r\times 2r}$. We have

$$\begin{aligned} &\sup_{\boldsymbol{T}\in\mathbb{R}^{m\times n}, \text{rank}(\boldsymbol{T})\le 2r, \|\boldsymbol{T}\|_F=1} \langle \boldsymbol{T}, \boldsymbol{Y} - \boldsymbol{X}\rangle \\ =& \sup_{\boldsymbol{T}\in\mathbb{R}^{m\times n}, \text{rank}(\boldsymbol{T})\le 2r, \|\boldsymbol{T}\|_F=1} \langle \boldsymbol{U}\Sigma, \boldsymbol{E}\boldsymbol{V}\rangle \\ =& \sup_{\boldsymbol{v}\in\mathbb{R}^{2nr}} \boldsymbol{v}^T\boldsymbol{e} \le C\sigma\sqrt{nr}, \end{aligned} \tag{B.36}$$

with probability $1 - \exp(-C_2 nr)$, where C, C_2 are two positive constants, the vectorization $\boldsymbol{e} = \text{Vec}(\boldsymbol{E}\boldsymbol{V}) \in \mathbb{R}^{2nr}$ has independent mean-zero sub-Gaussian entries with bounded variance σ^2 due to the orthogonality of $\boldsymbol{V}$, and the last inequality follows from (Rigollet and Hütter, 2015, Theorem 1.19).

Combining inequalities (B.35) and (B.36), we obtain the desired conclusion.

□

Proofs of Theorem 3.7 (Achievability) and Theorem 3.17

Proof of Theorem 3.7 (Achievability) and Theorem 3.17. The proofs of Theorem 3.7 (Achievability) and Theorem 3.17 share the same idea. We prove the contraction step by step. In each step, we show the specific procedures for the algorithm loss and address the MLE loss by stating the difference.

We consider dTBM (3.1) with symmetric mean tensor, parameters $(z, \mathcal{S}, \boldsymbol{\theta})$, fixed $r \ge 1, K \ge 2$, and i.i.d. noise. Let $(\hat{z}, \hat{\mathcal{S}}, \hat{\boldsymbol{\theta}})$ denote the MLE in (3.9), and $(z_k^{(0)}, \mathcal{S}^{(0)}, \boldsymbol{\theta}_k^{(0)})$ denote parameters related to the initialization. For the case $r = 1$, $\ell(z_k^{(t)}, z) = 0$ trivially for all $t \ge 0, k \in [k]$. Hence, we focus on the proof of the first mode clustering $z_1^{(t+1)}$ with $r \ge 2$; the extension for other modes can be obtained similarly. We drop the subscript k in the matricizations $\boldsymbol{\Theta}, \boldsymbol{M}_k, \boldsymbol{S}_k, \boldsymbol{X}_k$ and in estimates $z_k^{(0)}, z_k^{(t+1)}, z_k^{(t)}$ for ease of the notation. Without loss of generality, we assume that the variance $\sigma = 1$, and that the identity permutation minimizes

the initial misclustering error; i.e., $\pi^{(0)} = \arg\min_{\pi\in\Pi} \sum_{i\in[p]} \mathbb{1}\left\{z^{(0)}(i) \neq \pi \circ z(i)\right\}$ and $\pi^{(0)}(a) = a$ for all $a \in [r]$, and so for $\hat{z}$.

Step 1 (Notation and conditions). We first introduce additional notations and the necessary conditions used in the proof. We will verify that the conditions hold in our context under high probability in the last step of the proof.

Notation.

wide Projection. We use $\boldsymbol{I}_d$ to denote the identity matrix of dimension d. For a vector $\boldsymbol{v} \in \mathbb{R}^d$, let $\text{Proj}(\boldsymbol{v}) \in \mathbb{R}^{d\times d}$ denote the projection matrix to $\boldsymbol{v}$. Then, $\boldsymbol{I}_d - \text{Proj}(\boldsymbol{v})$ is the projection matrix to the orthogonal complement $\boldsymbol{v}^\perp$.

wiide We define normalized membership matrices

$$\boldsymbol{W} = \boldsymbol{M}\left(\text{diag}(\mathbf{1}_p^T\boldsymbol{M})\right)^{-1}, \boldsymbol{W}^{(t)} = \boldsymbol{M}^{(t)}\left(\text{diag}(\mathbf{1}_p^T\boldsymbol{M}^{(t)})\right)^{-1},$$

weighted normalized membership matrices

$$\begin{aligned}\boldsymbol{P} &= \boldsymbol{\Theta}\boldsymbol{M}(\text{diag}(\|\boldsymbol{\theta}_{z^{-1}(1)}\|^2, \ldots, \|\boldsymbol{\theta}_{z^{-1}(r)}\|^2))^{-1},\\ \hat{\boldsymbol{P}} &= \hat{\boldsymbol{\Theta}}\hat{\boldsymbol{M}}(\text{diag}(\|\hat{\boldsymbol{\theta}}_{z^{-1}(1)}\|^2, \ldots, \|\hat{\boldsymbol{\theta}}_{z^{-1}(r)}\|^2))^{-1},\end{aligned}$$

and the dual normalized and dual weighted normalized membership matrices

$$\begin{aligned}&\boldsymbol{V} = \boldsymbol{W}^{\otimes(K-1)}, \quad \boldsymbol{V}^{(t)} = \left(\boldsymbol{W}^{(t)}\right)^{\otimes(K-1)},\\ &\boldsymbol{Q} = \boldsymbol{P}^{\otimes K-1}, \quad \hat{\boldsymbol{Q}} = \hat{\boldsymbol{P}}^{\otimes K-1}.\end{aligned}$$

Also, let $\boldsymbol{B} = (\boldsymbol{\Theta}\boldsymbol{M})^{\otimes(K-1)}$, $\hat{\boldsymbol{B}} = (\hat{\boldsymbol{\Theta}}\hat{\boldsymbol{M}})^{\otimes(K-1)}$. By the definition, we have $\boldsymbol{B}^T\boldsymbol{Q} = \hat{\boldsymbol{B}}^T\hat{\boldsymbol{Q}} = \boldsymbol{I}_{r^{K-1}}$.

wiiide We use $\mathcal{S}^{(t)}$ to denote the estimator of $\mathcal{S}$ in the t-th iteration, $\hat{\mathcal{S}}$ for MLE, $\tilde{\mathcal{S}}$ to denote the oracle estimator of $\mathcal{S}$ given true assignment z, and $\bar{\mathcal{S}}$ for

weighted oracle estimator; i.e.,

$$\begin{aligned}
\mathcal{S}^{(t)} &= \mathcal{Y} \times_1 \left(\boldsymbol{W}^{(t)}\right)^T \times_2 \cdots \times_K \left(\boldsymbol{W}^{(t)}\right)^T, \\
\tilde{\mathcal{S}} &= \mathcal{Y} \times_1 \boldsymbol{W}^T \times_2 \cdots \times_K \boldsymbol{W}^T, \\
\hat{\mathcal{S}} &= \mathcal{Y} \times_1 \hat{\boldsymbol{P}}^T \times_2 \cdots \times_K \hat{\boldsymbol{P}}^T, \\
\bar{\mathcal{S}} &= \mathcal{Y} \times_1 \boldsymbol{P}^T \times_2 \cdots \times_K \boldsymbol{P}^T.
\end{aligned}$$

wivde We define the matricizations of tensors

$$\boldsymbol{S} = \text{Mat}(\mathcal{S}),\ \boldsymbol{Y} = \text{Mat}(\mathcal{Y}),\ \boldsymbol{X} = \text{Mat}(\mathcal{X}),\ \boldsymbol{E} = \text{Mat}(\mathcal{E}),$$

$$\boldsymbol{S}^{(t)} = \text{Mat}(\mathcal{S}^{(t)}),\ \hat{\boldsymbol{S}} = \text{Mat}(\hat{\mathcal{S}}),\ \tilde{\boldsymbol{S}} = \text{Mat}(\tilde{\mathcal{S}}),\ \bar{\boldsymbol{S}} = \text{Mat}(\bar{\mathcal{S}}).$$

wvde We define the extended core tensor on $K-1$ modes

$$\boldsymbol{A} = \boldsymbol{S}\boldsymbol{B}^T, \quad \bar{\boldsymbol{A}} = \bar{\boldsymbol{S}}\boldsymbol{B}^T, \quad \hat{\boldsymbol{A}} = \hat{\boldsymbol{S}}\hat{\boldsymbol{B}}^T.$$

By the assumption in parameter space (3.2), we have $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{X} = \boldsymbol{W}\boldsymbol{X}, \quad \hat{\boldsymbol{A}} = \hat{\boldsymbol{P}}\hat{\boldsymbol{X}} = \hat{\boldsymbol{W}}\hat{\boldsymbol{X}}$.

wvide We define the angle-based misclustering loss in the t-th iteration and loss for MLE

$$\begin{aligned}
L^{(t)} &= \frac{1}{p}\sum_{i\in[p]}\theta(i)\sum_{b\in[r]}\mathbb{1}\{z^{(t)}(i)=b\}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2, \\
L(\hat{z}) &= \frac{1}{p}\sum_{i\in[p]}\theta(i)^2\sum_{b\in[r]}\mathbb{1}\{\hat{z}(i)=b\}\|[\boldsymbol{A}_{z(i):}]^s - [\boldsymbol{A}_{b:}]^s\|^2.
\end{aligned}$$

We also define the loss for oracle and weighted oracle estimators

$$\begin{aligned}
\xi = \frac{1}{p}\sum_{i\in[p]}\theta(i)\sum_{b\in[r]}\mathbb{1}\Big\{ &\left\langle \boldsymbol{E}_{i:}\boldsymbol{V}, [\tilde{\boldsymbol{S}}_{z(i):}]^s - [\tilde{\boldsymbol{S}}_{b:}]^s \right\rangle \\
&\le -\frac{\theta(i)m}{4}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2\Big\} \\
&\cdot \|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2, \\
\xi' = \frac{1}{p}\sum_{i\in[p]}\theta(i)^2\sum_{b\in[r]}\mathbb{1}\Big\{ &\left\langle \boldsymbol{E}_{i:}, [\bar{\boldsymbol{A}}_{z(i):}]^s - [\bar{\boldsymbol{A}}_{b:}]^s \right\rangle \\
&\le -\frac{m'}{4}\sqrt{\frac{p^{K-1}}{r^{K-1}}}\|[\boldsymbol{A}_{z(i):}]^s - [\boldsymbol{A}_{b:}]^s\|_F^2\Big\} \\
&\cdot \|[\boldsymbol{A}_{z(i):}]^s - [\boldsymbol{A}_{b:}]^s\|^2.
\end{aligned}$$

where m and m' are some positive universal constants.

Then we introduce the necessary conditions in Condition 1.

Condition 1. (*Intermediate results*) *Let $\mathbb{O}_{p,r}$ denote the collection of all the p-by-r matrices with orthonormal columns. We have*

$$\begin{aligned}
&\|\boldsymbol{EV}\|_\sigma \lesssim \sqrt{\frac{r^{K-1}}{p^{K-1}}}\left(p^{1/2} + r^{(K-1)/2}\right), \quad \|\boldsymbol{EV}\|_F \lesssim \sqrt{\frac{r^{2(K-1)}}{p^{K-2}}}, \\
&\|\boldsymbol{W}_{a:}^T\boldsymbol{EV}\| \lesssim \frac{r^K}{p^{K/2}}, \text{ for all } a \in [r],
\end{aligned} \tag{B.37}$$

$$\sup_{\boldsymbol{U}_k\in\mathbb{O}_{p,r},k=2,\ldots,K}\|\boldsymbol{E}(\boldsymbol{U}_2\otimes\cdots\otimes\boldsymbol{U}_K)\|_\sigma \lesssim \left(\sqrt{r^{K-1}} + K\sqrt{pr}\right), \tag{B.38}$$

$$\sup_{\boldsymbol{U}_k\in\mathbb{O}_{p,r},k=2,\ldots,K}\|\boldsymbol{E}(\boldsymbol{U}_2\otimes\cdots\otimes\boldsymbol{U}_K)\|_F \lesssim \left(\sqrt{pr^{K-1}} + K\sqrt{pr}\right), \tag{B.39}$$

$$\xi \le \exp\left(-M\frac{\Delta_{\min}^2 p^{K-1}}{r^{K-1}}\right), \quad \xi' \lesssim \exp\left(-\frac{\Delta_{\min}^2 p^{K-1}}{r^{K-1}}\right), \tag{B.40}$$

$$L^{(t)} \leq \frac{\bar{C}}{\tilde{C}} \frac{\Delta_{\min}^2}{r \log p}, \quad \textit{for} \quad t = 0, 1, \ldots, T, \quad L(\hat{z}) \leq \frac{\bar{C}}{\tilde{C}} \frac{\Delta_{\min}^2}{r \log p}, \tag{B.41}$$

where M is a positive universal constant in inequality (B.59)*, $\bar{C}, \tilde{C}$ are positive universal constants in the proof of Theorem 3.12 and assumption SNR $\geq \tilde{C} p^{-K/2} \log p$, respectively. Further, inequality* (B.37) *holds by replacing $\boldsymbol{V}$ to $\boldsymbol{V}^{(t)}, \boldsymbol{Q}, \hat{\boldsymbol{Q}}$ and $\boldsymbol{W}_{:a}$ to $\boldsymbol{W}_{:a}^{(t),T}, \boldsymbol{P}_{:a}^T, \hat{\boldsymbol{P}}_{:a}^T$ when initialization condition* (B.41) *holds.*

Step 2 (Misclustering loss decomposition). Next, we derive the upper bound of $L^{(t+1)}$ for $t = 0, 1, \ldots, T-1$. By Sub-algorithm 2, we update the assignment in t-th iteration via

$$z^{(t+1)}(i) = \arg\min_{a \in [r]} \| [\boldsymbol{Y}_{i:} \boldsymbol{V}^{(t)}]^s - [\boldsymbol{S}_{a:}^{(t)}]^s \|^2,$$

following the facts that $\|\boldsymbol{a}^s - \boldsymbol{b}^s\|^2 = 1 - \cos(\boldsymbol{a}, \boldsymbol{b})$ for vectors $\boldsymbol{a}, \boldsymbol{b}$ of same dimension and $\mathrm{Mat}(\mathcal{Y}^{\mathrm{d}}) = \boldsymbol{Y}\boldsymbol{V}^{(t)}$ where $\mathcal{Y}^{\mathrm{d}}$ is the reduced tensor defined in Step 8 of Sub-algorithm 2. Then the event $z^{(t+1)}(i) = b$ implies

$$\| [\boldsymbol{Y}_{i:} \boldsymbol{V}^{(t)}]^s - [\boldsymbol{S}_{b:}^{(t)}]^s \|^2 \leq \| [\boldsymbol{Y}_{i:} \boldsymbol{V}^{(t)}]^s - [\boldsymbol{S}_{z(i):}^{(t)}]^s \|^2. \tag{B.42}$$

Note that the event (B.42) also holds for the degenerate entity i with $\|\boldsymbol{Y}_{i:} \boldsymbol{V}^{(t)}\| = 0$ due to the convention that $\boldsymbol{a}^s = \boldsymbol{0}$ if $\boldsymbol{a} = \boldsymbol{0}$. Arranging the terms in (B.42) yields the decomposition

$$\begin{aligned} &2 \left\langle \boldsymbol{E}_{i:} \boldsymbol{V}, [\tilde{\boldsymbol{S}}_{z(i):}]^s - [\tilde{\boldsymbol{S}}_{b:}]^s \right\rangle \\ &\quad \leq \|\boldsymbol{X}_{i:} \boldsymbol{V}^{(t)}\| \left(-\| [\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s \|^2 + G_{ib}^{(t)} + H_{ib}^{(t)} \right) + F_{ib}^{(t)}, \end{aligned}$$

where

$$
\begin{aligned}
F_{ib}^{(t)} &= 2\left\langle \boldsymbol{E}_{i:}\boldsymbol{V}^{(t)}, \left([\tilde{\boldsymbol{S}}_{z(i):}]^s - [\boldsymbol{S}_{z(i):}^{(t)}]^s\right) - \left([\tilde{\boldsymbol{S}}_{b:}]^s - [\boldsymbol{S}_{b:}^{(t)}]^s\right)\right\rangle \\
&\quad + 2\left\langle \boldsymbol{E}_{i:}\left(\boldsymbol{V} - \boldsymbol{V}^{(t)}\right), [\tilde{\boldsymbol{S}}_{z(i):}]^s - [\tilde{\boldsymbol{S}}_{b:}]^s\right\rangle, \qquad \text{(B.43)} \\
G_{ib}^{(t)} &= \Big(\|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{S}_{z(i):}^{(t)}]^s\|^2 \\
&\qquad\qquad - \|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:z(i)}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^2\Big) \\
&\quad - \Big(\|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{S}_{b:}^{(t)}]^s\|^2 \\
&\qquad\qquad - \|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^2\Big), \qquad \text{(B.44)} \\
H_{ib}^{(t)} &= \|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:z(i)}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^2 - \|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s \\
&\quad - [\boldsymbol{W}_{:b}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^2 + \|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2. \qquad \text{(B.45)}
\end{aligned}
$$

Therefore, the event $\mathbb{1}\left\{z^{(t+1)}(i) = b\right\}$ can be upper bounded as

$$
\begin{aligned}
&\mathbb{1}\left\{z^{(t+1)}(i) = b\right\} \\
&\le \mathbb{1}\Bigg\{z^{(t+1)}(i) = b, \left\langle \boldsymbol{E}_{j:}\boldsymbol{V}, [\tilde{\boldsymbol{S}}_{z(i):}]^s - [\tilde{\boldsymbol{S}}_{b:}]^s\right\rangle \\
&\qquad\qquad \le -\frac{1}{4}\|\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}\|\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2\Bigg\} \\
&\quad + \mathbb{1}\Bigg\{z^{(t+1)}(i) = b, \frac{1}{2}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2 \\
&\qquad\qquad \le \|\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}\|^{-1}F_{ib}^{(t)} + G_{ib}^{(t)} + H_{ib}^{(t)}\Bigg\}. \qquad \text{(B.46)}
\end{aligned}
$$

Note that

$$
\begin{aligned}
\|\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}\| &= \theta(i)\|\boldsymbol{S}_{i:}(\boldsymbol{\Theta}\boldsymbol{M})^{\otimes(K-1),T}\boldsymbol{W}^{(t),\otimes^{K-1}}\| \\
&\ge \theta(i)\|\boldsymbol{S}_{z(i):}\|\lambda_r^{K-1}(\boldsymbol{\Theta}\boldsymbol{M})\lambda_r^{K-1}(\boldsymbol{W}^{(t)}) \\
&\ge \theta(i)m, \qquad \text{(B.47)}
\end{aligned}
$$

where the first inequality follows from the property of eigenvalues; the last inequal-

ity follows from Lemma B.6, Lemma B.10, and assumption that $\min_{a\in[r]}\|\boldsymbol{S}_{z(i):}\| \geq c_3 > 0$; and $m > 0$ is a positive constant related to c_3. Plugging the lower bound of $\|\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}\|$ (B.47) into the inequality (B.46) gives

$$\mathbb{1}\left\{z^{(t+1)}(i) = b\right\} \leq A_{ib} + B_{ib}, \tag{B.48}$$

where

$$\begin{aligned}
A_{ib} = \mathbb{1}\Bigg\{ & z^{(t+1)}(i) = b, \left\langle \boldsymbol{E}_{i:}\boldsymbol{V}, [\tilde{\boldsymbol{S}}_{z(i):}]^s - [\tilde{\boldsymbol{S}}_{b:}]^s \right\rangle \\
& \leq -\frac{\theta(i)m}{4}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2 \Bigg\}, \\
B_{ib} = \mathbb{1}\Bigg\{ & z^{(t+1)}(i) = b, \frac{1}{2}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2 \\
& \leq (\theta(i)m)^{-1}F_{ib}^{(t)} + G_{ib}^{(t)} + H_{ib}^{(t)} \Bigg\}.
\end{aligned}$$

Taking the weighted summation of (B.48) over $i \in [p]$ yields

$$L^{(t+1)} \leq \xi + \frac{1}{p}\sum_{i\in[p]}\sum_{b\in[r]/z(i)}\zeta_{ib}^{(t)},$$

where ξ is the oracle loss such that

$$\xi = \frac{1}{p}\sum_{i\in[p]}\theta(i)\sum_{b\in[r]/z(i)}A_{ib}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2. \tag{B.49}$$

Similarly to ξ in (B.49), we define

$$\zeta_{ib}^{(t)} = \theta(i)B_{ib}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2.$$

Now, we show the decomposition for MLE loss.
By the definition of Gaussian MLE, the estimator $\hat{\boldsymbol{\theta}}$ satisfies $\hat{\theta}(i) = \left\langle \boldsymbol{Y}_{i:}, \hat{\boldsymbol{A}}_{\hat{z}(i):} \right\rangle / \|\hat{\boldsymbol{A}}_{\hat{z}(i):}\|_F^2$

for all $i \in [p]$. Hence, we have

$$\hat{z}(i) = \underset{a\in[r_1]}{\arg\min} \|[\boldsymbol{Y}_{i:}]^s - [\hat{\boldsymbol{A}}_{a:}]^s\|_F^2,$$

and the decomposition

$$L(\hat{z}) \le \xi' + \frac{1}{p}\sum_{i\in[p]}\sum_{b\in[r]/z(i)} \zeta'_{ib},$$

where $\zeta'_{ib} = \theta(i)^2 B'_{ib}\|[\boldsymbol{A}_{z(i):}]^s - [\boldsymbol{A}_{b:}]^s\|^2$ and

$$\begin{aligned}
A'_{ib} = \mathbb{1}\Bigg\{ & \hat{z}(i) = b, \left\langle \boldsymbol{E}_{i:}, [\bar{\boldsymbol{A}}_{z(i):}]^s - [\bar{\boldsymbol{A}}_{b:}]^s \right\rangle \\
& \le -\frac{m'}{4}\sqrt{\frac{p^{K-1}}{r^{K-1}}}\|[\boldsymbol{A}_{z(i):}]^s - [\boldsymbol{A}_{b:}]^s\|_F^2 \Bigg\}, \\
B'_{ib} = \mathbb{1}\Bigg\{ & \hat{z}(i) = b, -\frac{1}{2}\|[\boldsymbol{A}_{z(i):}]^s - [\boldsymbol{A}_{b:}]^s\|_F^2 \\
& \le \sqrt{\frac{r^{K-1}}{(m')^2 p^{K-1}}}\hat{F}_{ib} + \hat{G}_{ib} + \hat{H}_{ib} \Bigg\}
\end{aligned}$$

with terms

$$\begin{aligned}
\hat{F}_{ib} &= 2\left\langle \boldsymbol{E}_{i:}, ([\bar{\boldsymbol{A}}_{z(i):}]^s - [\hat{\boldsymbol{A}}_{a:}]^s) - ([\bar{\boldsymbol{A}}_{b:}]^s - [\hat{\boldsymbol{A}}_{b:}]^s) \right\rangle, \\
\hat{G}_{ib} &= \left(\|\boldsymbol{X}^s_{i:} - [\hat{\boldsymbol{A}}_{z(i):}]^s\|_F^2 - \|\boldsymbol{X}^s_{i:} - [\boldsymbol{P}^T_{:z(i)}\boldsymbol{Y}\hat{\boldsymbol{Q}}\hat{\boldsymbol{B}}^T]^s\|_F^2\right) \\
&\quad - \left(\|\boldsymbol{X}^s_{i:} - [\hat{\boldsymbol{A}}_{b:}]^s\|_F^2 - \|\boldsymbol{X}^s_{i:} - [\boldsymbol{P}^T_{:b}\boldsymbol{Y}\hat{\boldsymbol{Q}}\hat{\boldsymbol{B}}^T]^s\|_F^2\right), \\
\hat{H}_{ib} &= \|\boldsymbol{X}^s_{i:} - [\boldsymbol{P}^T_{:z(i)}\boldsymbol{Y}\hat{\boldsymbol{Q}}\hat{\boldsymbol{B}}^T]^s\|_F^2 - \|\boldsymbol{X}^s_{i:} - [\boldsymbol{P}^T_{:b}\boldsymbol{Y}\hat{\boldsymbol{Q}}\hat{\boldsymbol{B}}^T]^s\|_F^2 \\
&\quad + \|\boldsymbol{A}^s_{z(i):} - \boldsymbol{A}^s_{b:}\|_F^2.
\end{aligned}$$

Step 3 (Derivation of contraction inequality). In this step we derive the upper bound of ζ_{ib} and obtain the contraction inequality (3.24). We show the analysis in the following one-column box for a better presentation.

Choose the constant $\tilde{C}$ in the condition SNR $\ge \tilde{C}p^{-K/2}\log p$ that satisfies the

condition of Lemma B.11, inequalities (B.73), and (B.77). Note that

$$\zeta_{ib}^{(t)} = \theta(i)\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2 \mathbb{1}\left\{z^{(t+1)}(i) = b, \frac{1}{2}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2 \leq (\theta(i)m)^{-1}F_{ib}^{(t)} + G_{ib}^{(t)} + H_{ib}^{(t)}\right\}$$

$$\leq \theta(i)\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2 \mathbb{1}\left\{z^{(t+1)}(i) = b, \frac{1}{4}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2 \leq (\theta(i)m)^{-1}F_{ib}^{(t)} + G_{ib}^{(t)}\right\}$$

$$\leq 64\mathbb{1}\left\{z^{(t+1)}(i) = b\right\}\left(\frac{(F_{ib}^{(t)})^2}{cm^2\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2} + \frac{\theta(i)(G_{ib}^{(t)})^2}{\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2}\right)$$

where the first inequality follows from the inequality (B.64) in Lemma B.11, and the last inequality follows from the assumption that $\min_{i\in[p]}\theta(i) \geq c > 0$. Following (Han et al., 2022a, Step 4, Proof of Theorem 2) and Lemma B.11, we have

$$\frac{1}{p}\sum_{i\in[p]}\sum_{b\in[r]/z(i)}\mathbb{1}\left\{z^{(t+1)}(i) = b\right\}\frac{(F_{ib}^{(t)})^2}{cm^2\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2} \leq \frac{C_0\bar{C}}{cm^2\tilde{C}^2}L^{(t)},$$

for a positive universal constant C and

$$\begin{aligned}
&\frac{1}{p}\sum_{i\in[p]}\sum_{b\in[r]/z(i)}\mathbb{1}\left\{z^{(t+1)}(i) = b\right\}\frac{\theta(i)(G_{ib}^{(t)})^2}{\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2} \\
&\leq \frac{1}{512}\frac{1}{p}\sum_{i\in[p]}\theta(i)\sum_{b\in[r]/z(i)}\mathbb{1}\left\{z^{(t+1)}(i) = b\right\}(\Delta_{\min}^2 + L^{(t)}) \\
&\leq \frac{1}{512}(L^{(t+1)} + L^{(t)}),
\end{aligned}$$

where the last inequality follows from the definition of $L^{(t)}$ and the constraint of $\boldsymbol{\theta}$ in parameter space (3.2). For $\tilde{C}$ also satisfies

$$\frac{C_0\bar{C}}{cm^2\tilde{C}^2} \leq \frac{1}{512}, \tag{B.50}$$

we have

$$\frac{1}{p}\sum_{i\in[p]}\sum_{b\in[r]/z(i)}\zeta_{ib}^{(t)} \le \frac{1}{8}L^{(t+1)} + \frac{1}{4}L^{(t)}. \tag{B.51}$$

Plugging the inequality (B.51) into the decomposition (B.49), we obtain the contraction inequality

$$L^{(t+1)} \le \frac{3}{2}\xi + \frac{1}{2}L^{(t)}, \tag{B.52}$$

where $\frac{1}{2}$ is the contraction parameter.

Therefore, with $\tilde{C}$ satisfying inequalities (B.50), (B.73) and (B.77), we obtain the conclusion in Theorem 3.17 via inequality (B.52) combining the inequality (B.40) in Condition 1 and Lemma 3.11.

We also have the contraction inequality for MLE.

Following the same derivation of (B.52) with the upper bound of $\hat{F}_{ib}, \hat{G}_{ib}, \hat{H}_{ib}$ in Lemma B.12, we also have

$$L(\hat{z}) \le \frac{3}{2}\xi' + \frac{1}{2}L(\hat{z}),$$

which indicates the conclusion $\ell(\hat{z}, z) \lesssim \Delta_{\min}^2 \exp\left(-\frac{p^{K-1}}{r^{K-1}}\Delta_{\min}^2\right)$.

Step 4 (Verification of Condition 1). Last, we verify the Condition 1 under high probability to finish the proof. Note that the inequalities (B.37), (B.38), and (B.39) describe the property of the sub-Gaussian noise tensor $\mathcal{E}$, and the readers can find the proof directly in (Han et al., 2022a, Step 5, Proof of Theorem 2). The initial condition (B.41) for MLE is satisfied by Lemma B.13. Here, we include only the verification of inequalities (B.40) and (B.41) for algorithm estimators.

Now, we verify the oracle loss condition (B.40). Recall the definition of ξ,

$$\begin{aligned}\xi = \frac{1}{p}\sum_{i\in[p]}\theta(i)\sum_{b\in[r]}\mathbb{1}\Big\{ &\left\langle \boldsymbol{E}_{i:}\boldsymbol{V}, [\tilde{\boldsymbol{S}}_{z(i):}]^s - [\tilde{\boldsymbol{S}}_{b:}]^s \right\rangle \\ &\le -\frac{\theta(i)m}{4}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2\Big\} \\ &\cdot \|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2.\end{aligned}$$

Let $e_i = \boldsymbol{E}_{i:}\boldsymbol{V}$ denote the aggregated noise vector for all $i \in [p]$, and e_i's are inde-

pendent zero-mean sub-Gaussian vector in $\mathbb{R}^{r^{K-1}}$. The entries in e_i are independent zero-mean sub-Gaussian variables with sub-Gaussian norm upper bounded by $m_1\sqrt{r^{K-1}/p^{K-1}}$ with some positive constant m_1. We have the probability inequality

$$\begin{aligned}\mathbb{P}\left(\left\langle e_i, [\tilde{\boldsymbol{S}}_{z(i):}]^s - [\tilde{\boldsymbol{S}}_{b:}]^s\right\rangle \leq -\frac{\theta(i)m}{4}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2\right)\\ \leq P_1 + P_2 + P_3,\end{aligned}$$

where

$$\begin{aligned}P_1 &= \mathbb{P}\left(\left\langle e_i, [\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\right\rangle \leq -\frac{\theta(i)m}{8}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2\right),\\ P_2 &= \mathbb{P}\left(\left\langle e_i, [\tilde{\boldsymbol{S}}_{z(i):}]^s - [\boldsymbol{S}_{z(i):}]^s\right\rangle \leq -\frac{\theta(i)m}{16}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2\right),\\ P_3 &= \mathbb{P}\left(\left\langle e_i, [\boldsymbol{S}_{b:}]^s - [\tilde{\boldsymbol{S}}_{b:}]^s\right\rangle \leq -\frac{\theta(i)m}{16}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2\right).\end{aligned}$$

For P_1, notice that the inner product $\left\langle e_j, \boldsymbol{S}^s_{z(j):} - \boldsymbol{S}^s_{b:}\right\rangle$ is a sub-Gaussian variable with sub-Gaussian norm bounded by $m_2\sqrt{r^{K-1}/p^{K-1}}\|\boldsymbol{S}^s_{z(i):} - \boldsymbol{S}^s_{b:}\|$ with some positive constant m_2. Then, by Chernoff bound, we have

$$P_1 \lesssim \exp\left(-\frac{p^{K-1}}{r^{K-1}}\|[\boldsymbol{S}_{z(j):}]^s - [\boldsymbol{S}_{b:}]^s\|^2\right). \tag{B.53}$$

For P_2 and P_3, we only need to derive the upper bound of P_2 due to the symmetry. By the law of total probability, we have

$$P_2 \leq P_{21} + P_{22}, \tag{B.54}$$

where with some positive constant $t > 0$,

$$P_{21} = \mathbb{P}\left(t \leq \|[\tilde{\boldsymbol{S}}_{z(i):}]^s - [\boldsymbol{S}_{z(i):}]^s\|\right),$$
$$P_{22} = \mathbb{P}\Bigg(\left\langle e_i, [\tilde{\boldsymbol{S}}_{z(i):}]^s - [\boldsymbol{S}_{z(i):}]^s\right\rangle \leq -\frac{\theta(i)m}{16} \cdot \|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2 \Bigg| \|[\tilde{\boldsymbol{S}}_{z(i):}]^s - [\boldsymbol{S}_{z(i):}]^s\| < t\Bigg).$$

For P_{21}, note that the term $\boldsymbol{W}_{:z(i)}^T \boldsymbol{E}\boldsymbol{V} = \frac{\sum_{j\neq i, j\in[p]} \mathbb{1}\{z(j)=z(i)\}e_j}{\sum_{j\in[p]} \mathbb{1}\{z(j)=z(i)\}}$ is a sub-Gaussian vector with sub-Gaussian norm bounded by $m_3\sqrt{r^K/p^K}$ with some positive constant m_3. This implies

$$\begin{aligned} P_{21} &\leq \mathbb{P}\left(t\|\boldsymbol{S}_{z(i):}\| \leq \|\tilde{\boldsymbol{S}}_{z(i):} - \boldsymbol{S}_{z(i):}\|\right) \\ &\leq \mathbb{P}\left(c_3 t \leq \|\boldsymbol{W}_{:z(i)}^T \boldsymbol{E}\boldsymbol{V}\|\right) \\ &\lesssim \exp\left(-\frac{p^K t^2}{r^K}\right), \end{aligned} \tag{B.55}$$

where the first inequality follows from the basic inequality in Lemma B.3, the second inequality follows from the assumption that $\min_{a\in[r]}\|\boldsymbol{S}_{z(i):}\| \geq c_3 > 0$ in (3.2), and the last inequality follows from the Bernstein inequality.

For P_{22}, the inner product $\left\langle e_i, [\tilde{\boldsymbol{S}}_{z(i):}]^s - [\boldsymbol{S}_{z(i):}]^s\right\rangle$ is also a sub-Gaussian variable with sub-Gaussian norm $m_4\sqrt{r^{K-1}/p^{K-1}}t$, conditioned on $\|[\tilde{\boldsymbol{S}}_{z(i):}]^s - [\boldsymbol{S}_{z(i):}]^s\| < t$ with some positive constant m_4. Then, by Chernoff bound, we have

$$P_{22} \lesssim \exp\left(-\frac{p^{K-1}}{r^{K-1}t^2}\|[\boldsymbol{S}_{z(j):}]^s - [\boldsymbol{S}_{b:}]^s\|^4\right). \tag{B.56}$$

We take $t = \|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|$ in P_{21} and P_{22}, and plug the inequalities (B.55) and (B.56) into to the upper bound for P_2 in (B.54). We obtain that

$$P_2 \lesssim \exp\left(-\frac{p^{K-1}}{r^{K-1}}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2\right). \tag{B.57}$$

Combining the upper bounds (B.53) and (B.57) gives

$$\begin{aligned}
&\mathbb{P}\left(\left\langle e_i, [\tilde{\boldsymbol{S}}_{z(i):}]^s - [\tilde{\boldsymbol{S}}_{b:}]^s \right\rangle \le -\frac{\theta(i)m}{4}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2\right) \\
&\lesssim \exp\left(-\frac{p^{K-1}}{r^{K-1}}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2\right). \qquad \text{(B.58)}
\end{aligned}$$

Hence, we have

$$\begin{aligned}
\mathbb{E}\xi &= \frac{1}{p}\sum_{i\in[p]}\theta(i)\sum_{b\in[r]}\mathbb{P}\Bigg\{\left\langle \boldsymbol{E}_{i:}\boldsymbol{V}, [\tilde{\boldsymbol{S}}_{z(i):}]^s - [\tilde{\boldsymbol{S}}_{b:}]^s\right\rangle \\
&\qquad \le -\frac{\theta(i)m}{4}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2\Bigg\}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2 \\
&\lesssim \frac{1}{p}\sum_{i\in[p]}\theta(i)\max_{i\in[p],b\in[r]}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2 \\
&\qquad\qquad \cdot \exp\left(-\frac{p^{K-1}}{r^{K-1}}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2\right) \\
&\le \exp\left(-M\frac{p^{K-1}}{r^{K-1}}\Delta_{\min}^2\right), \qquad \text{(B.59)}
\end{aligned}$$

where M is a positive constant, the first inequality follows from the constraint that $\sum_{i\in[p]}\theta(i) = p$, and the last inequality follows from (B.58).

By Markov's inequality, we have

$$\begin{aligned}
&\mathbb{P}\left(\xi \lesssim \mathbb{E}\xi + \exp\left(-\frac{Mp^{K-1}}{2r^{K-1}}\Delta_{\min}^2\right)\right) \\
&\qquad\qquad \ge 1 - C\exp\left(-\frac{Mp^{K-1}}{2r^{K-1}}\Delta_{\min}^2\right),
\end{aligned}$$

and thus the condition (B.40) holds with probability at least $1 - C\exp\left(-\frac{Mp^{K-1}}{2r^{K-1}}\Delta_{\min}^2\right)$ for some constant $C > 0$.

The initialization condition for MLE also holds.

For ξ', notice that $\langle \boldsymbol{E}_i, \boldsymbol{A}^s_{a:} - \boldsymbol{A}^s_{b:}\rangle$ is a sub-Gaussian vector with variance bounded

by $\|\boldsymbol{A}^s_{a:} - \boldsymbol{A}^s_{b:}\|^2$ and

$$\begin{aligned}\mathbb{P}\left(t \le \|[\bar{\boldsymbol{A}}_{a:}]^s - \boldsymbol{A}^s_{a:}\|\right) &\le \left(t \le \|[\boldsymbol{P}^T_{:a}\boldsymbol{Y}\boldsymbol{Q}]^s - [\boldsymbol{P}^T_{:a}\boldsymbol{X}\boldsymbol{Q}]^s\|\right) \\ &\le \mathbb{P}(t \min_{a\in[r]}\|\boldsymbol{S}_{a:}\| \le \|\boldsymbol{P}^T_{:a}\boldsymbol{E}\boldsymbol{Q}\|) \\ &\lesssim \exp\left(-\frac{p^K t^2}{r^K}\right),\end{aligned}$$

where the first inequality follows from the property in later inequality (B.80). We also have

$$\xi' \lesssim \left(-\frac{p^{K-1}}{r^{K-1}}\Delta^2_{\min}\right).$$

Finally, we verify the bounded loss condition (B.41) for algorithm estimator by induction. With output $z^{(0)}$ from Sub-algorithm 2 and the assumption SNR $\ge \tilde{C}p^{-K/2}\log p$, by Theorem 3.12, we have

$$L^{(0)} \le \frac{\bar{C}\Delta^2_{\min}}{\tilde{C}r\log p}, \quad \text{when } p \text{ is large enough.}$$

Therefore, the condition (B.41) holds for $t = 0$. Assume that the condition (B.41) also holds for all $t \le t_0$. Then, by the decomposition (B.52), we have

$$\begin{aligned}L^{(t_0+1)} &\le \frac{3}{2}\xi + \frac{1}{2}L^{(t_0)} \\ &\le \exp\left(-M\frac{p^{K-1}}{r^{K-1}}\Delta^2_{\min}\right) + \frac{\Delta^2_{\min}}{r\log p} \\ &\le \frac{\bar{C}}{\tilde{C}}\frac{\Delta^2_{\min}}{r\log p},\end{aligned}$$

where the second inequality follows from the condition (B.40) and the last inequality follows from the assumption that $\Delta^2_{\min} \gtrsim p^{-K/2}\log p$. Thus, the condition (B.41) holds for $t_0 + 1$, and the condition (B.41) is proved by induction. □

Useful Lemmas for the Proof of Theorem 3.17

Lemma B.10 (Singular-value property of membership matrices)**.** *Under the setup of Theorem 3.17, suppose that the condition* (B.41) *holds. Then, for all* $a \in [r]$*, we have* $|\left(z^{(t)}\right)^{-1}(a)| \asymp p/r$*. Moreover, we have*

$$\begin{aligned} &\lambda_r(\boldsymbol{M}) \asymp \|\boldsymbol{M}\|_\sigma \asymp \sqrt{p/r},\ \lambda_r(\boldsymbol{W}) \asymp \|\boldsymbol{W}\|_\sigma \asymp \sqrt{r/p}, \\ &\lambda_r(\boldsymbol{P}) \asymp \|\boldsymbol{P}\|_\sigma \asymp \min_{a\in[r]} \|\boldsymbol{\theta}_{z^{-1}(a)}\|^{-1} \lesssim \sqrt{r/p}. \end{aligned} \tag{B.60}$$

The inequalities (B.60) *also hold by replacing* $\boldsymbol{M}$ *and* $\boldsymbol{W}$ *to* $\boldsymbol{M}^{(t)}$ *and* $\boldsymbol{W}^{(t)}$ *respectively. Further, we have*

$$\lambda_r(\boldsymbol{W}\boldsymbol{W}^T) \asymp \left\|\boldsymbol{W}\boldsymbol{W}^T\right\|_\sigma \asymp r/p, \tag{B.61}$$

which is also true for $\boldsymbol{W}^{(t)}\boldsymbol{W}^{(t),T}$*.*

Proof of Lemma B.10. The proof for the inequality (B.60) for $\boldsymbol{M}, \boldsymbol{W}$ can be found in (Han et al., 2022a, Proof of Lemma 4). The inequalities for $\boldsymbol{P}$ follows the same derivation with balance assumption on $\boldsymbol{\theta}$ and $\min_{i\in[p]} \theta(i) \geq c$.

For inequality (B.61), note that for all $k \in [r]$,

$$\begin{aligned} \lambda_k(\boldsymbol{W}\boldsymbol{W}^T) &= \sqrt{\text{eigen}_k(\boldsymbol{W}\boldsymbol{W}^T\boldsymbol{W}\boldsymbol{W}^T)} \\ &\asymp \sqrt{\frac{r}{p}\text{eigen}_k(\boldsymbol{W}\boldsymbol{W}^T)} \\ &= \sqrt{\frac{r}{p}\lambda_k^2(\boldsymbol{W})} \asymp \frac{r}{p}, \end{aligned}$$

where $\text{eigen}_k(\boldsymbol{A})$ denotes the k-th largest eigenvalue of the square matrix $\boldsymbol{A}$, the first inequality follows the fact that $\boldsymbol{W}^T\boldsymbol{W}$ is a diagonal matrix with elements of order r/p, and the second equation follows from the definition of singular value. □

Lemma B.11 (Upper bound for $F_{ib}^{(t)}, G_{ib}^{(t)}$ and $H_{ib}^{(t)}$)**.** *Under the Condition 1 and the setup of Theorem 3.17 with fixed* $r \geq 2$*, assume the constant* $\tilde{C}$ *in the condition SNR* $\geq \tilde{C}p^{-K/2}\log p$ *is large enough to satisfy the inequalities* (B.73) *and* (B.77)*. As* $p \to \infty$*, we have*

$$\max_{i\in[p]}\max_{b\neq z(i)}\frac{\left(F_{ib}^{(t)}\right)^2}{\|[\boldsymbol{S}_{z(i)\cdot}]^s-[\boldsymbol{S}_{b\cdot}]^s\|^2}$$
$$\lesssim\frac{rL^{(t)}}{\Delta_{\min}^2}\|\boldsymbol{E}_{i\cdot}\boldsymbol{V}\|^2+\left(1+\frac{rL^{(t)}}{\Delta_{\min}^2}\right)\|\boldsymbol{E}_{i\cdot}(\boldsymbol{V}-\boldsymbol{V}^{(t)})\|^2, \tag{B.62}$$

$$\max_{i\in[p]}\max_{b\neq z(i)}\frac{\left(G_{ib}^{(t)}\right)^2}{\|[\boldsymbol{S}_{z(i)\cdot}]^s-[\boldsymbol{S}_{b\cdot}]^s\|^2}\leq\frac{1}{512}\left(\Delta_{\min}^2+L^{(t)}\right), \tag{B.63}$$

$$\max_{i\in[p]}\max_{b\neq z(i)}\frac{\left|H_{ib}^{(t)}\right|}{\|[\boldsymbol{S}_{z(i)\cdot}]^s-[\boldsymbol{S}_{b\cdot}]^s\|^2}\leq\frac{1}{4}. \tag{B.64}$$

Similarly, when the SNR $\geq\tilde{C}p^{-(K-1)}\log p$ *with a large constant* $\tilde{C}$*, we have*

$$\max_{i\in[p]}\max_{b\neq z(i)}\frac{\left(\hat{F}_{ib}\right)^2}{\|[\boldsymbol{A}_{z(i)\cdot}]^s-[\boldsymbol{A}_{b\cdot}]^s\|^2}\lesssim p^{K-1}\frac{rL(\hat{z})}{\Delta_{\min}^2}$$

$$\max_{i\in[p]}\max_{b\neq z(i)}\frac{\left(\hat{G}_{ib}\right)^2}{\|[\boldsymbol{A}_{z(i)\cdot}]^s-[\boldsymbol{A}_{b\cdot}]^s\|^2}\leq\frac{1}{512}\left(\Delta_{\min}^2+L(\hat{z})\right),$$

$$\max_{i\in[p]}\max_{b\neq z(i)}\frac{\left|\hat{H}_{ib}\right|}{\|[\boldsymbol{A}_{z(i)\cdot}]^s-[\boldsymbol{A}_{b\cdot}]^s\|^2}\leq\frac{1}{4}.$$

Proof of Lemma B.11. We prove the the first three inequalities in Lemma B.11 separately.

wide Upper bound for $F_{ib}^{(t)}$, i.e., inequality (B.62). Recall the definition of $F_{ib}^{(t)}$,

$$\begin{aligned}F_{ib}^{(t)}&=2\left\langle\boldsymbol{E}_{i\cdot}\boldsymbol{V}^{(t)},\left([\tilde{\boldsymbol{S}}_{z(i)\cdot}]^s-[\boldsymbol{S}_{z(i)\cdot}^{(t)}]^s\right)-\left([\tilde{\boldsymbol{S}}_{b\cdot}]^s-[\boldsymbol{S}_{b\cdot}^{(t)}]^s\right)\right\rangle\\&\quad+2\left\langle\boldsymbol{E}_{i\cdot}(\boldsymbol{V}-\boldsymbol{V}^{(t)}),[\tilde{\boldsymbol{S}}_{z(i)\cdot}]^s-[\tilde{\boldsymbol{S}}_{b\cdot}]^s\right\rangle.\end{aligned}$$

By Cauchy-Schwartz inequality, we have

$$
\begin{aligned}
&\left(F_{ib}^{(t)}\right)^2 \\
&\le 8\left(\left\langle \boldsymbol{E}_{i:}\boldsymbol{V}^{(t)}, \left([\tilde{\boldsymbol{S}}_{z(i):}]^s - [\boldsymbol{S}_{z(i):}^{(t)}]^s\right) - \left([\tilde{\boldsymbol{S}}_{b:}]^s - [\boldsymbol{S}_{b:}^{(t)}]^s\right)\right\rangle\right)^2 \\
&\qquad + 8\left(\left\langle \boldsymbol{E}_{i:}(\boldsymbol{V} - \boldsymbol{V}^{(t)}), [\tilde{\boldsymbol{S}}_{z(i):}]^s - [\tilde{\boldsymbol{S}}_{b:}]^s\right\rangle\right)^2 \\
&\le 8\left(\|\boldsymbol{E}_{i:}\boldsymbol{V}\|^2 + \|\boldsymbol{E}_{i:}(\boldsymbol{V} - \boldsymbol{V}^{(t)})\|^2\right) \max_{a\in[r]^s} \|[\tilde{\boldsymbol{S}}_{a:}]^s - [\boldsymbol{S}_{a:}^{(t)}]^s\| \\
&\qquad + \|\boldsymbol{E}_{i:}(\boldsymbol{V} - \boldsymbol{V}^{(t)})\|^2 \|[\tilde{\boldsymbol{S}}_{z(i):}]^s - [\tilde{\boldsymbol{S}}_{b:}]^s\|. \qquad \text{(B.65)}
\end{aligned}
$$

Note that for all $a \in [r]$,

$$
\begin{aligned}
\|[\tilde{\boldsymbol{S}}_{a:}]^s - [\boldsymbol{S}_{a:}^{(t)}]^s\|^2 &= \|[\boldsymbol{W}_{:a}^T\boldsymbol{Y}\boldsymbol{V}]^s - [\boldsymbol{W}_{:a}^{(t),T}\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^2 \\
&\le 2\|[\boldsymbol{W}_{:a}^T\boldsymbol{Y}\boldsymbol{V}]^s - [\boldsymbol{W}_{:a}^{(t),T}\boldsymbol{Y}\boldsymbol{V}]^s\|^2 \\
&\quad + 2\|[\boldsymbol{W}_{:a}^{(t),T}\boldsymbol{Y}\boldsymbol{V}]^s - [\boldsymbol{W}_{:a}^{(t),T}\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^2 \\
&\lesssim \frac{r^2(L^{(t)})^2}{\Delta_{\min}^2} + \frac{rr^{2K} + pr^{K+2}}{p^K}\frac{L^{(t)}}{\Delta_{\min}^2} \\
&\lesssim rL^{(t)} + \frac{rr^{2K} + pr^{K+2}}{p^K}\frac{L^{(t)}}{\Delta_{\min}^2} \\
&\lesssim rL^{(t)}, \qquad \text{(B.66)}
\end{aligned}
$$

where the second inequality follows from the inequalities (B.83) and (B.84) in Lemma B.12, the third inequality follows from the condition (B.41) in Condition 1, and the last inequality follows from the assumption that $\Delta_{\min}^2 \ge \tilde{C}p^{-K/2}\log p$.

Note that

$$
\begin{aligned}
&\|[\tilde{\boldsymbol{S}}_{z(i):}]^s - [\tilde{\boldsymbol{S}}_{b:}]^s\|^2 \\
&= \|[\tilde{\boldsymbol{S}}_{z(i):}]^s - [\boldsymbol{S}_{z(i):}]^s + [\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s + [\boldsymbol{S}_{b:}]^s - [\tilde{\boldsymbol{S}}_{b:}]^s\|^2 \\
&\lesssim \|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2 + \max_{a\in[r]} \|[\boldsymbol{S}_{a:}]^s - [\tilde{\boldsymbol{S}}_{a:}]^s\|^2 \\
&\lesssim \|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2 + \max_{a\in[r]} \frac{1}{\|\boldsymbol{S}_{a:}\|^2} \|\boldsymbol{W}_{:a}^T \boldsymbol{E}\boldsymbol{V}\|^2 \\
&\lesssim \|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2, \qquad \text{(B.67)}
\end{aligned}
$$

where the second inequality follows from Lemma B.3, and the last inequality follows from the assumptions on $\|\boldsymbol{S}_{a:}\|$ in the parameter space (3.2), the inequality (B.37) in Condition 1 and the assumption $\Delta_{\min}^2 \gtrsim p^{-K/2}\log p$.

Therefore, we finish the proof of inequality (B.62) by plugging the inequalities (B.66) and (B.67) into the upper bound (B.65).

wiide Upper bound for $G_{ib}^{(t)}$, i.e., inequality (B.63). By definition of $G_{ib}^{(t)}$, we rearrange terms and obtain

$$
\begin{aligned}
G_{ib}^{(t)} = &\Big(\|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{S}_{z(i):}^{(t)}]^s\|^2 \\
&\qquad - \|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:z(i)}^T \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^2 \Big) \\
&- \Big(\|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{S}_{b:}^{(t)}]^s\|^2 \\
&\qquad - \|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^T \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^2 \Big) \\
= &\, 2\Big\langle [\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s, \Big([\boldsymbol{W}_{:z(i)}^T \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{S}_{z(i):}^{(t)}]^s \Big) \\
&\qquad - \Big([\boldsymbol{W}_{:b}^T \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{S}_{b:}^{(t)}]^s \Big) \Big\rangle \\
= &\, G_1 + G_2 - G_3, \qquad \text{(B.68)}
\end{aligned}
$$

where

$$
\begin{aligned}
G_1 &= \|[\boldsymbol{W}_{:z(i)}^T \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{S}_{z(i):}^{(t)}]^s\|^2 - \|[\boldsymbol{W}_{:b}^T \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{S}_{b:}^{(t)}]^s\|^2, \\
G_2 &= 2\left\langle [\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:z(i)}^T \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s, [\boldsymbol{W}_{:z(i)}^T \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{S}_{z(i):}^{(t)}]^s \right\rangle, \\
G_3 &= 2\left\langle [\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^T \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s, [\boldsymbol{W}_{:b}^T \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{S}_{b:}^{(t)}]^s \right\rangle.
\end{aligned}
$$

For G_1, we have

$$
\begin{aligned}
|G_1|^2 &\leq \Big| \|[\boldsymbol{W}_{:z(i)}^T \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{S}_{z(i):}^{(t)}]^s\|^2 \\
&\qquad\qquad - \|[\boldsymbol{W}_{:b}^T \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{S}_{b:}^{(t)}]^s\|^2 \Big|^2 \\
&\leq \max_{a\in[r]} \|[\boldsymbol{W}_{:a}^T \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:a}^{(t),T} \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^4 \\
&\leq C^4 \frac{r^4}{\Delta_{\min}^4} (L^{(t)})^4 + \frac{r^2 r^{4K} + p^2 r^{2K+4}}{p^{2K}} \frac{(L^{(t)})^2}{\Delta_{\min}^4} \\
&\leq C^4 \frac{\bar{C}}{\tilde{C}^3} \left(\Delta_{\min}^4 + \Delta_{\min}^2 L^{(t)} \right), \qquad \text{(B.69)}
\end{aligned}
$$

where the third inequality follows from the inequality (B.85) in Lemma B.12 and the last inequality follows from the assumption that $\Delta_{\min}^2 \geq \tilde{C} p^{-K/2} \log p$ and inequality (B.41) in Condition 1.

For G_2, noticing that $[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s = [\boldsymbol{W}_{z(i):}^T\boldsymbol{X}\boldsymbol{V}^{(t)}]^s$, we have

$$\begin{aligned}
|G_2|^2 &\le 2\|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:z(i)}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^2 \\
&\qquad \cdot \|[\boldsymbol{W}_{:z(i)}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{S}_{z(i):}^{(t)}]^s\|^2 \\
&\le \frac{2}{\|\boldsymbol{W}_{z(i):}^T\boldsymbol{X}\boldsymbol{V}^{(t)}\|^2}\max_{a\in[r]}\|\boldsymbol{W}_{:a}^T\boldsymbol{E}\boldsymbol{V}^{(t)}\|^2 \\
&\qquad \cdot \max_{a\in[r]}\|[\boldsymbol{W}_{:a}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:a}^{(t),T}\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^2 \\
&\le C'\frac{r^{2K-1}+Kpr^{K+1}}{p^K} \\
&\qquad \cdot \left(\frac{r^2}{\Delta_{\min}^2}(L^{(t)})^2 + \frac{rr^{2K}+pr^{K+2}}{p^K}\frac{L^{(t)}}{\Delta_{\min}^2}\right) \\
&\le \frac{C'}{\tilde{C}^2}\Delta_{\min}^2 L^{(t)},
\end{aligned} \tag{B.70}$$

where C' is a positive universal constant, the second inequality follows from Lemma B.3, the third inequality follows from the inequality (B.38) in Condition 1, the inequalities (B.85) and (B.104) in the proof of Lemma B.12, and the last inequality follows from the assumption $\Delta_{\min}^2 \ge \tilde{C}p^{-K/2}\log p$ and inequality (B.41) in Condition 1.

For G_3, note that by triangle inequality

$$\begin{aligned}
&\|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^T\boldsymbol{X}\boldsymbol{V}^{(t)}]^s\|^2 \\
&\le \|\boldsymbol{S}_{z(i):}^s - \boldsymbol{S}_{b:}^s\|^2 + 2\max_{a\in[r]}\|[\boldsymbol{W}_{:a}^T\boldsymbol{X}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:a}^T\boldsymbol{X}\boldsymbol{V}]^s\|^2 \\
&\le \|\boldsymbol{S}_{z(i):}^s - \boldsymbol{S}_{b:}^s\|^2 + C\frac{r^2(L^{(t)})^2}{\Delta_{\min}^2},
\end{aligned} \tag{B.71}$$

where the last inequality follows from the inequality (B.103) in the proof of Lemma B.12 and C is a positive constant. Then we have

$$
\begin{aligned}
|G_3|^2 &\le 2\|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^2 \\
&\qquad \cdot \max_{a\in[r]}\|[\boldsymbol{W}_{:a}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:a}^{(t),T}\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^2 \\
&\le 2\Big(\|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^T\boldsymbol{X}\boldsymbol{V}^{(t)}]^s\|^2 \\
&\qquad + \|[\boldsymbol{W}_{:b}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^T\boldsymbol{X}\boldsymbol{V}^{(t)}]^s\|^2\Big) \\
&\quad \cdot \max_{a\in[r]}\left\|[\boldsymbol{W}_{:a}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:a}^{(t),T}\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\right\|^2 \\
&\le C^2\left(\|\boldsymbol{S}_{z(i):}^s - \boldsymbol{S}_{b:}^s\|^2 + C\frac{r^2(L^{(t)})^2}{\Delta_{\min}^2}\right) \\
&\quad \cdot \left(\frac{r^2(L^{(t)})^2}{\Delta_{\min}^2} + \frac{rr^{2K}+pr^{K+2}}{p^K}\frac{L^{(t)}}{\Delta_{\min}^2}\right) + \frac{C'}{\tilde{C}^2}\Delta_{\min}^2 L^{(t)} \\
&\le \frac{C^2\bar{C}^2}{\tilde{C}}\|\boldsymbol{S}_{z(i):}^s - \boldsymbol{S}_{b:}^s\|^2(\Delta_{\min}^2 + L^{(t)}) \\
&\quad + \frac{C^3C'\bar{C}^2}{\tilde{C}^2}\left(\Delta_{\min}^4 + \Delta_{\min}^2 L^{(t)}\right),
\end{aligned} \tag{B.72}
$$

where the third inequality follows from the same procedure to derive (B.69) and (B.70), and the last inequality follows from the assumption $\Delta_{\min}^2 \ge \tilde{C}p^{-K/2}\log p$ and inequality (B.41) in Condition 1.

Choose the $\tilde{C}$ such that

$$
3\left(C^4\frac{\bar{C}}{\tilde{C}^3} + \frac{C'}{\tilde{C}^2} + \frac{C^2\bar{C}^2}{\tilde{C}} + \frac{C^3C'\bar{C}^2}{\tilde{C}^2}\right) \le \frac{1}{512}. \tag{B.73}
$$

Then, we finish the proof of inequality (B.63) by plugging the inequalities (B.69), (B.70), and (B.72) into the upper bound (B.68).

wiiide Upper bound for $H_{ib}^{(t)}$, i.e., the inequality (B.64). By definition of H_{ib}, we rearrange terms and obtain

$$
\begin{aligned}
H_{ib} &= \|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:z(i)}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^2 \\
&\quad - \|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^2 + \|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2 \\
&= \|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:z(i)}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^2 \\
&\quad + \Big(\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2 - \|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^T\boldsymbol{X}\boldsymbol{V}^{(t)}]^s\|\Big) \\
&\quad - \Big(\|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\| \\
&\qquad\qquad - \|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^T\boldsymbol{X}\boldsymbol{V}^{(t)}]^s\|\Big) \\
&= H_1 + H_2 + H_3,
\end{aligned}
$$

where

$$
\begin{aligned}
H_1 &= \|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:z(i)}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^2 \\
&\qquad\qquad - \|[\boldsymbol{W}_{:b}^T\boldsymbol{X}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\|^2, \\
H_2 &= \|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2 - \|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^T\boldsymbol{X}\boldsymbol{V}^{(t)}]^s\|^2, \\
H_3 &= 2\Big\langle[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^T\boldsymbol{X}\boldsymbol{V}^{(t)}]^s, \\
&\qquad\qquad [\boldsymbol{W}_{:b}^T\boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^T\boldsymbol{X}\boldsymbol{V}^{(t)}]^s\Big\rangle.
\end{aligned}
$$

For H_1, we have

$$
\begin{aligned}
|H_1| &\le \frac{4\max_{a\in[r]}\|\boldsymbol{W}_{:a}^T\boldsymbol{E}\boldsymbol{V}^{(t)}\|^2}{\|\boldsymbol{W}_{z(i):}^T\boldsymbol{X}\boldsymbol{V}^{(t)}\|^2} \\
&\le \frac{r^{2K-1} + Kpr^{K+1}}{p^K} \\
&\le \tilde{C}^{-2}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{b:}]^s\|^2, \qquad \text{(B.74)}
\end{aligned}
$$

following the derivation of G_2 in inequality (B.70) and the assumption that $\Delta_{\min}^2 \ge \tilde{C}p^{-K/2}\log p$.

For H_2, by the inequality (B.71), we have

$$\begin{aligned}|H_2| &\lesssim 2\max_{a\in[r]}\|[\boldsymbol{W}_{:a}^T\boldsymbol{X}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:a}^T\boldsymbol{X}\boldsymbol{V}]^s\|^2 \\ &\lesssim \frac{r^2(L^{(t)})^2}{\Delta_{\min}^2} \\ &\leq C\frac{\bar{C}^2}{\tilde{C}^2}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{a:}]^s\|^2, \end{aligned} \tag{B.75}$$

where the last inequality follows from the condition (B.41) in Condition 1.

For H_3, by Cauchy-Schwartz inequality, we have

$$\begin{aligned}|H_3| &\lesssim \|[\boldsymbol{X}_{i:}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^T\boldsymbol{X}\boldsymbol{V}^{(t)}]^s\||H_1|^{1/2} \\ &\leq 2\tilde{C}^{-1}\|[\boldsymbol{S}_{z(i):}]^s - [\boldsymbol{S}_{a:}]^s\|^2, \end{aligned} \tag{B.76}$$

following the inequalities (B.71) and (B.74).

Choose $\tilde{C}$ such that

$$\tilde{C}^{-2} + C\frac{\bar{C}^2}{\tilde{C}^2} + \tilde{C}^{-1} \leq \frac{1}{4}. \tag{B.77}$$

Therefore, we finish the proof of inequality (B.64) combining inequalities (B.74), (B.75), and (B.76).

Next, we show the upper bounds for $\hat{F}_{ib}, \hat{G}_{ib}$ and $\hat{H}_{ib}$.
By Lemma 3.4, we have

$$\|\boldsymbol{S}_{a:}^s - \boldsymbol{S}_{b:}^s\| = (1+o(1))\|\boldsymbol{A}_{a:}^s - \boldsymbol{A}_{b:}^s\|.$$

Also, notice that the matrix product of $\boldsymbol{B}^T$ corresponds to the padding operation in Lemma B.5, and the padding weights are balanced such that $\|\boldsymbol{v}\boldsymbol{B}\| = (1 + o(1))\max_a\|\boldsymbol{\theta}_{z^{-1}(a)}\|^{(K-1)/2}\|\boldsymbol{v}\|$ for all $\boldsymbol{v} \in \mathbb{R}^{r(K-1)}$. For two vectors $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{R}^{r^{K-1}}$, we have

$$\|\boldsymbol{v}_1^s - \boldsymbol{v}_2^s\| = (1+o(1))\|[\boldsymbol{v}_1\boldsymbol{B}^T]^s - [\boldsymbol{v}_2\boldsymbol{B}^T]^s\|. \tag{B.78}$$

The equation (B.78) also holds for $\hat{\boldsymbol{B}}^T$.

Note that for all $i \in [p]$ we have

$$\begin{aligned}\|\boldsymbol{A}_{i:}\hat{\boldsymbol{Q}}\| &= \|\boldsymbol{S}_{z(i:)}\boldsymbol{B}^T\hat{\boldsymbol{Q}}\| \\ &= \|\boldsymbol{S}_{z(i:)}\hat{\boldsymbol{D}}^{\otimes(K-1)}\| \\ &= (1+o(1))\|\boldsymbol{S}_{z(i:)}\| \\ &= (1+o(1))\max_a \|\boldsymbol{\theta}_{z^{-1}(a)}\|^{-(K-1)/2}\|\boldsymbol{A}_{i:}\|, \end{aligned} \tag{B.79}$$

where the third inequality follows from the singular property of MLE confusion matrix (B.110) and the last inequality follows from the fact that $\boldsymbol{A}_i = \boldsymbol{S}_{z(i:)}\boldsymbol{B}^T$ and Lemma B.10. Above equation indicates that $\boldsymbol{A}_{i:}$ is the span space of the singular values as $p \to \infty$. Also, notice that the row space of $\boldsymbol{P}_{:a}^T\boldsymbol{Y}\hat{\boldsymbol{Q}}\hat{\boldsymbol{B}}^T$ is equal to the column space of $\hat{\boldsymbol{Q}}$, and $\boldsymbol{A}_{i:} \neq \boldsymbol{P}_{:a}^T\boldsymbol{Y}\hat{\boldsymbol{Q}}\hat{\boldsymbol{B}}^T$ in noisy case.

Hence, for all $a \in [r]$, we have

$$\begin{aligned}&\|[\boldsymbol{X}_i\hat{\boldsymbol{Q}}]^s - [\boldsymbol{P}_{:a}^T\boldsymbol{Y}\hat{\boldsymbol{Q}}]^s\| \\ &= \left\| \frac{\boldsymbol{A}_{z(i:)}\hat{\boldsymbol{Q}}}{\|\boldsymbol{A}_{z(i:)}\hat{\boldsymbol{Q}}\|} - \frac{\boldsymbol{P}_{:a}^T\boldsymbol{Y}\hat{\boldsymbol{Q}}}{\|\boldsymbol{P}_{:a}^T\boldsymbol{Y}\hat{\boldsymbol{Q}}\|} \right\| \\ &= (1+o(1))\left\| \frac{\boldsymbol{A}_{z(i:)}}{\|\boldsymbol{A}_{z(i:)}\|} - \frac{\boldsymbol{P}_{:a}^T\boldsymbol{Y}\hat{\boldsymbol{Q}}\hat{\boldsymbol{B}}^T}{\|\boldsymbol{P}_{:a}^T\boldsymbol{Y}\hat{\boldsymbol{Q}}\hat{\boldsymbol{B}}^T\|} \right\| \\ &= (1+o(1))\|[\boldsymbol{X}_i]^s - [\boldsymbol{P}_{:a}^T\boldsymbol{Y}\hat{\boldsymbol{Q}}\hat{\boldsymbol{B}}^T]^s\| \end{aligned} \tag{B.80}$$

where the second equation follows from (B.79), $\|\boldsymbol{P}_{:a}^T\boldsymbol{Y}\hat{\boldsymbol{Q}}\hat{\boldsymbol{B}}^T\| = (1+o(1))\max_a \|\boldsymbol{\theta}_{z^{-1}(a)}\|^{(K-1)/2}\|\boldsymbol{P}_{:a}^T\boldsymbol{Y}\hat{\boldsymbol{Q}}\|$, and singular property of $\hat{\boldsymbol{B}}^T$. Similar result holds after replacing $\boldsymbol{P}_{:a}^T\boldsymbol{Y}\hat{\boldsymbol{Q}}$ by $\boldsymbol{P}_{:a}^T\boldsymbol{Y}\boldsymbol{Q}$ or $\boldsymbol{P}_{:a}^T\boldsymbol{Y}\hat{\boldsymbol{Q}}$.

We are now ready to show the upper bounds for $\hat{F}_{ib}, \hat{G}_{ib}$ and $\hat{H}_{ib}$.

For $\hat{F}_{ib}$, we have

$$\begin{aligned}(\hat{F}_{ib})^2 &\leq \|\boldsymbol{E}_{i:}\|^2 \|[\bar{\boldsymbol{A}}_{a:}]^s - [\hat{\boldsymbol{A}}_{a:}]^s\|^2 \\ &\leq \|\boldsymbol{E}_{i:}\|^2 \Big[\|[\bar{\boldsymbol{S}}_{a:}\boldsymbol{B}^T]^s - [\bar{\boldsymbol{S}}_{a:}\hat{\boldsymbol{B}}^T]^s\| \\ &\qquad + \|[\bar{\boldsymbol{S}}_{a:}\hat{\boldsymbol{B}}^T]^s - [\hat{\boldsymbol{S}}_{a:}\hat{\boldsymbol{B}}^T]^s\| \Big]^2 \\ &\lesssim \|\boldsymbol{E}_{i:}\|^2 \left[\|[\bar{\boldsymbol{S}}_{a:}\boldsymbol{B}^T\hat{\boldsymbol{Q}}]^s - [\bar{\boldsymbol{S}}_{a:}]^s\| + \|[\bar{\boldsymbol{S}}_{a:}]^s - [\hat{\boldsymbol{S}}_{a:}]^s\| \right]^2.\end{aligned}$$

Following similar derivations in inequalities (B.66), (B.67), and the upper bound for J_1 in the proof of Lemma B.12, respectively, we have

$$\|[\bar{\boldsymbol{S}}_{a:}]^s - [\hat{\boldsymbol{S}}_{a:}]^s\| \lesssim rL(\hat{z}), \quad \|[\bar{\boldsymbol{S}}_{a:}]^s - [\bar{\boldsymbol{S}}_{b:}]^s\| \lesssim \|\boldsymbol{S}^s_{a:} - \boldsymbol{S}^s_{b:}\|^2,$$

and

$$\|[\bar{\boldsymbol{S}}_{a:}\boldsymbol{B}^T\hat{\boldsymbol{Q}}]^s - [\bar{\boldsymbol{S}}_{a:}]^s\| \lesssim L(\hat{z}).$$

We then obtain the upper bound for $\hat{F}_{ib}$ by noticing that $\|\boldsymbol{E}_i\|^2 \lesssim p^{K-1}$.

For $\hat{G}_{ib}$ and $\hat{H}_{ib}$, by the property (B.80), we have

$$\begin{aligned}&(1+o(1))\hat{G}_{ib} \\ &= \left(\|[\boldsymbol{X}_{i:}\hat{\boldsymbol{Q}}]^s - [\hat{\boldsymbol{S}}_{a:}]^s\|_F^2 - \|[\boldsymbol{X}_{i:}\hat{\boldsymbol{Q}}]^s - [\boldsymbol{P}_{:a}^T\boldsymbol{Y}\hat{\boldsymbol{Q}}]^s\|_F^2 \right) \\ &\quad - \left(\|[\boldsymbol{X}_{i:}\hat{\boldsymbol{Q}}]^s - [\hat{\boldsymbol{S}}_{b:}]^s\|_F^2 - \|[\boldsymbol{X}_{i:}\hat{\boldsymbol{Q}}]^s - [\boldsymbol{P}_{:b}^T\boldsymbol{Y}\hat{\boldsymbol{Q}}]^s\|_F^2 \right), \\ &(1+o(1))\hat{H}_{ib} \\ &= \|[\boldsymbol{X}_{i:}\hat{\boldsymbol{Q}}]^s - [\boldsymbol{P}_{:a}^T\boldsymbol{Y}\hat{\boldsymbol{Q}}]^s\|_F^2 - \|[\boldsymbol{X}_{i:}\hat{\boldsymbol{Q}}]^s - [\boldsymbol{P}_{:b}^T\boldsymbol{Y}\hat{\boldsymbol{Q}}]^s\|_F^2 \\ &\quad + \|\boldsymbol{A}^s_{a:} - \boldsymbol{A}^s_{b:}\|_F^2.\end{aligned}$$

We obtain the upper bounds following the proof for inequalities (B.63) and (B.64).

□

Lemma B.12 (Relationship between misclustering loss and intermediate parame-

ters). *Under the Condition 1 and the setup of Theorem 3.17 with fixed $r \geq 2$, as $p \to \infty$, we have*

$$\|\boldsymbol{V} - \boldsymbol{V}^{(t)}\|_\sigma \lesssim \sqrt{\frac{r^{K-1}}{p^{K-1}}} \frac{r}{\Delta_{\min}^2} L^{(t)}, \tag{B.81}$$

$$\|\boldsymbol{E}(\boldsymbol{V} - \boldsymbol{V}^{(t)})\|_\sigma \lesssim \sqrt{\frac{r^{K-1}(pr^{K-1} + pr)}{p^{K-1}}} \frac{r}{\Delta_{\min}^2} L^{(t)}, \tag{B.82}$$

$$\begin{aligned} &\max_{b\in[r]} \|[\boldsymbol{W}_{:b}^T \boldsymbol{Y}\boldsymbol{V}]^s - [\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{Y}\boldsymbol{V}]^s\| \\ &\qquad \leq C \left(\frac{rL^{(t)}}{\Delta_{\min}} + \sqrt{\frac{r^{2K} + pr^{K+1}}{p^K}} \frac{\sqrt{L^{(t)}}}{\Delta_{\min}} \right), \end{aligned} \tag{B.83}$$

$$\begin{aligned} &\max_{b\in[r]} \|[\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{Y}\boldsymbol{V}]^s - [\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\| \\ &\qquad \leq C \left(\sqrt{\frac{rr^{2K} + pr^{K+2}}{p^K}} \frac{\sqrt{L^{(t)}}}{\Delta_{\min}} + \frac{rL^{(t)}}{\Delta_{\min}} \right), \end{aligned} \tag{B.84}$$

$$\begin{aligned} &\max_{b\in[r]} \|[\boldsymbol{W}_{:b}^T \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\| \\ &\qquad \leq C \left(\frac{rL^{(t)}}{\Delta_{\min}} + \sqrt{\frac{rr^{2K} + pr^{K+2}}{p^K}} \frac{\sqrt{L^{(t)}}}{\Delta_{\min}} \right), \end{aligned} \tag{B.85}$$

for some positive universal constant C. In addition, the inequality (B.84) *also holds by replacing $\boldsymbol{W}_{:b}^{(t)}$ to $\boldsymbol{W}_{:b}$. Further, the above inequalities holds after replacing $\boldsymbol{W}$ to $\boldsymbol{P}$, $\boldsymbol{V}$ to $\boldsymbol{Q}$, and $L^{(t)}$ to $L(\hat{z})$.*

Proof of Lemma B.12. We follow and use several intermediate conclusions in (Han et al., 2022a, Proof of Lemma 5). We prove each inequality separately.

wide Inequality (B.81). By (Han et al., 2022a, Proof of Lemma 5), we have

$$\|\boldsymbol{V} - \boldsymbol{V}^{(t)}\|_\sigma \lesssim \sqrt{\frac{r^{K-1}}{p^{K-1}}} r\ell^{(t)}.$$

Then, we complete the proof of inequality (B.81) by applying Lemma 3.11 to the above inequality.

wiide Inequality (B.82). By (Han et al., 2022a, Proof of Lemma 5), we have

$$\|\boldsymbol{E}(\boldsymbol{V}-\boldsymbol{V}^{(t)})\|_{\sigma} \lesssim \sqrt{\frac{r^{K-1}(pr^{K-1}+pr)}{p^{K-1}}} r\ell^{(t)}.$$

Also, we complete the proof of inequality (B.81) by applying Lemma 3.11 to the above inequality.

wiiide Inequality (B.83). We upper bound the desired quantity by triangle inequality,

$$\|[\boldsymbol{W}_{:b}^{T}\boldsymbol{Y}\boldsymbol{V}]^{s} - [\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{Y}\boldsymbol{V}]^{s}\| \leq I_1 + I_2 + I_3,$$

where

$$\begin{aligned}
I_1 &= \left\| \frac{\boldsymbol{W}_{:b}^{T}\boldsymbol{Y}\boldsymbol{V}}{\|\boldsymbol{W}_{:b}^{T}\boldsymbol{X}\boldsymbol{V}\|} - \frac{\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{Y}\boldsymbol{V}}{\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}\|} \right\|, \\
I_2 &= \left\| \left(\frac{1}{\|\boldsymbol{W}_{:b}^{T}\boldsymbol{Y}\boldsymbol{V}\|} - \frac{1}{\|\boldsymbol{W}_{:b}^{T}\boldsymbol{X}\boldsymbol{V}\|} \right) \boldsymbol{W}_{:b}^{T}\boldsymbol{Y}\boldsymbol{V} \right\|, \\
I_3 &= \left\| \left(\frac{1}{\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{Y}\boldsymbol{V}\|} - \frac{1}{\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}\|} \right) \boldsymbol{W}_{:b}^{(t),T}\boldsymbol{Y}\boldsymbol{V} \right\|.
\end{aligned}$$

Next, we upper bound the quantities I_1, I_2, I_3 separately.

For I_1, we further bound I_1 by triangle inequality,

$$I_1 \leq I_{11} + I_{12},$$

where

$$I_{11} = \left\| \frac{\boldsymbol{W}_{:b}^{T}\boldsymbol{X}\boldsymbol{V}}{\|\boldsymbol{W}_{:b}^{T}\boldsymbol{X}\boldsymbol{V}\|} - \frac{\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}}{\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}\|} \right\|,$$

and

$$I_{12} = \left\| \frac{\boldsymbol{W}_{:b}^{T}\boldsymbol{E}\boldsymbol{V}}{\|\boldsymbol{W}_{:b}^{T}\boldsymbol{X}\boldsymbol{V}\|} - \frac{\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{E}\boldsymbol{V}}{\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}\|} \right\|.$$

We first consider I_{11}. Define the confusion matrix $\boldsymbol{D} = \boldsymbol{M}^T \boldsymbol{\Theta}^T \boldsymbol{W}^{(t)} = [\![D_{ab}]\!] \in \mathbb{R}^{r \times r}$ where

$$D_{ab} = \frac{\sum_{i \in [p]} \theta(i) \mathbb{1}\left\{z(i) = a, z^{(t)}(i) = b\right\}}{\sum_{i \in [p]} \mathbb{1}\left\{z^{(t)}(i) = b\right\}}, \text{ for all } a, b \in [r].$$

By Lemma B.10, we have $\sum_{i \in [p]} \mathbb{1}\left\{z^{(t)}(i) = b\right\} \gtrsim p/r$. Then, we have

$$\sum_{a \neq b, a, b \in [r]} D_{ab} \lesssim \frac{r}{p} \sum_{i:\ z^{(t)}(i) \neq z(i)} \theta(i) \lesssim \frac{L^{(t)}}{\Delta_{\min}^2} \lesssim \frac{1}{\log p}, \tag{B.86}$$

and for all $b \in [r]$,

$$\begin{aligned} D_{bb} &= \frac{\sum_{i \in [p]} \theta(i) \mathbb{1}\left\{z(i) = z^{(t)}(i) = b\right\}}{\sum_{i \in [p]} \mathbb{1}\left\{z^{(t)}(i) = b\right\}} \\ &\geq \frac{c(\sum_{i \in [p]} \mathbb{1}\left\{z^{(t)}(i) = b\right\} - p\ell^{(t)})}{\sum_{i \in [p]} \mathbb{1}\left\{z^{(t)}(i) = b\right\}} \\ &\gtrsim 1 - \frac{1}{\log p}, \end{aligned} \tag{B.87}$$

under the inequality (B.41) in Condition 1. By the definition of $\boldsymbol{W}, \boldsymbol{W}^{(t)}, \boldsymbol{V}$, we have

$$\frac{\boldsymbol{W}_{:b}^T \boldsymbol{X} \boldsymbol{V}}{\|\boldsymbol{W}_{:b}^T \boldsymbol{X} \boldsymbol{V}\|} = [\boldsymbol{S}_{b:}]^s,$$

and

$$\frac{\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X} \boldsymbol{V}}{\|\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X} \boldsymbol{V}\|} = [D_{bb} \boldsymbol{S}_{b:} + \sum_{a \neq b, a \in [r]} D_{ab} \boldsymbol{S}_{a:}]^s.$$

Let α denote the angle between $\boldsymbol{S}_{b:}$ and $D_{bb} \boldsymbol{S}_{b:} + \sum_{a \neq b, a \in [r]} D_{ab} \boldsymbol{S}_{a:}$. To roughly

estimate the range of α, we consider the inner product

$$\begin{aligned}
&\left\langle \boldsymbol{S}_{b:}, D_{bb}\boldsymbol{S}_{b:} + \sum_{a\neq b, a\in[r]} D_{ab}\boldsymbol{S}_{a:} \right\rangle \\
&= D_{bb}\left\|\boldsymbol{S}_{b:}\right\|^2 + \sum_{a\neq b} D_{ab}\left\langle \boldsymbol{S}_{b:}, \boldsymbol{S}_{a:}\right\rangle \\
&\geq D_{bb}\left\|\boldsymbol{S}_{b:}\right\|^2 - \sum_{a\neq b, a\in[r]} D_{ab}\left\|\boldsymbol{S}_{b:}\right\| \max_{a\in[r]}\left\|\boldsymbol{S}_{a:}\right\| \\
&\geq C,
\end{aligned}$$

where C is a positive constant, and the last inequality holds when p is large enough following the constraint of $\|\boldsymbol{S}_{b:}\|$ in parameter space (3.2) and the bounds of $\boldsymbol{D}$ in (B.86) and (B.87).

The positive inner product between $\boldsymbol{S}_{b:}$ and $D_{bb}\boldsymbol{S}_{b:} + \sum_{a\neq b, a\in[r]} D_{ab}\boldsymbol{S}_{a:}$ indicates $\alpha \in [0, \pi/2)$, and thus $2\sin\frac{\alpha}{2} \leq \sqrt{2}\sin\alpha$. Then, by the geometry property of trigonometric function, we have

$$\begin{aligned}
&\left\|[D_{bb}\boldsymbol{S}_{b:} + \sum_{a\neq b, a\in[r]} D_{ab}\boldsymbol{S}_{a:}]\sin\alpha\right\| \\
&= \left\|(\boldsymbol{I}_d - \text{Proj}(\boldsymbol{S}_{b:}))\sum_{a\neq b, a\in[r]} D_{ab}\boldsymbol{S}_{a:}\right\| \\
&\leq \sum_{a\neq b, a\in[r]} D_{ab}\left\|(\boldsymbol{I}_d - \text{Proj}(\boldsymbol{S}_{b:}))\boldsymbol{S}_{a:}\right\| \\
&= \sum_{a\neq b, a\in[r]} D_{ab}\left\|\boldsymbol{S}_{a:}\sin(\boldsymbol{S}_{b:}, \boldsymbol{S}_{a:})\right\| \\
&\leq \sum_{a\neq b, a\in[r]} D_{ab}\left\|\boldsymbol{S}_{a:}\right\|\left\|\boldsymbol{S}^s_{b:} - \boldsymbol{S}^s_{a:}\right\|, \qquad \text{(B.88)}
\end{aligned}$$

where the first inequality follows from the triangle inequality, and the last inequality follows from Lemma B.3. Note that with bounds (B.86) and

(B.87), when p is large enough, we have

$$\begin{aligned}\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}\| &= \|D_{bb}\boldsymbol{S}_{b:} + \sum_{a\neq b, a\in[r]} D_{ab}\boldsymbol{S}_{a:}\| \\ &\geq D_{bb}\,\|\boldsymbol{S}_{b:}\| - \sum_{a\neq b, a\in[r]} D_{ab}\,\|\boldsymbol{S}_{a:}\| \\ &\geq C_1, \end{aligned} \tag{B.89}$$

for some positive constant C_1. Notice that $I_{11} = \sqrt{1-\cos\alpha} = 2\sin\frac{\alpha}{2}$. Therefore, we obtain

$$\begin{aligned} I_{11} &\leq \sqrt{2}\sin\alpha \\ &= \frac{\|[D_{bb}\boldsymbol{S}_{b:} + \sum_{a\neq b, a\in[r]} D_{ab}\boldsymbol{S}_{a:}]\sin\alpha\|}{\|D_{bb}\boldsymbol{S}_{b:} + \sum_{a\neq b, a\in[r]} D_{ab}\boldsymbol{S}_{a:}\|} \\ &\leq \frac{1}{C_1}\sum_{a\neq b, a\in[r]} D_{ab}\,\|\boldsymbol{S}_{a:}\|\,\|\boldsymbol{S}_{b:}^s - \boldsymbol{S}_{a:}^s\| \\ &\lesssim \frac{r}{p}\sum_{i\in[p]}\theta(i)\sum_{b\in[r]}\mathbb{1}\left\{z^{(t)}(i)=b\right\}\|\boldsymbol{S}_{b:}^s - \boldsymbol{S}_{a:}^s\| \\ &\leq \frac{rL^{(t)}}{\Delta_{\min}}, \end{aligned} \tag{B.90}$$

where the second inequality follows from (B.88) and (B.89), and the last two inequalities follow by the definition of D_a and $L^{(t)}$, and the constraint of $\|\boldsymbol{S}_{b:}\|$ in parameter space (3.2).

We now consider I_{12}. By triangle inequality, we have

$$\begin{aligned} I_{12} \leq{}& \frac{1}{\|\boldsymbol{W}_{:b}^T\boldsymbol{X}\boldsymbol{V}\|}\|(\boldsymbol{W}_{:b}^T - \boldsymbol{W}_{:b}^{(t),T})\boldsymbol{E}\boldsymbol{V}\| \\ &+ \frac{\|(\boldsymbol{W}_{:b}^T - \boldsymbol{W}_{:b}^{(t),T})\boldsymbol{X}\boldsymbol{V}\|}{\|\boldsymbol{W}_{:b}^T\boldsymbol{X}\boldsymbol{V}\|\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}\|}\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{E}\boldsymbol{V}\|. \end{aligned}$$

By (Han et al., 2022a, Proof of Lemma 5), we have

$$\|(\boldsymbol{W}_{:b}^T - \boldsymbol{W}_{:b}^{(t),T})\boldsymbol{E}\boldsymbol{V}\| \lesssim \sqrt{\frac{r^{2K} + pr^{K+1}}{p^K}} \frac{\sqrt{L^{(t)}}}{\Delta_{\min}}. \tag{B.91}$$

Notice that

$$\begin{aligned}
\|(\boldsymbol{W}_{:b}^T - \boldsymbol{W}_{:b}^{(t),T})\boldsymbol{X}\boldsymbol{V}\| &\leq \|\boldsymbol{W}_{:b}^T - \boldsymbol{W}_{:b}^{(t),T}\| \, \|\boldsymbol{X}\boldsymbol{V}\|_F \\
&\lesssim \frac{r^{3/2} L^{(t)}}{\sqrt{p}\Delta_{\min}^2} \|\boldsymbol{S}\| \|\boldsymbol{\Theta}\boldsymbol{M}\|_\sigma \\
&\lesssim \frac{\sqrt{rL^{(t)}}}{\Delta_{\min}},
\end{aligned} \tag{B.92}$$

where the second inequality follows from (Han et al., 2022a, Inequality (121), Proof of Lemma 5) and the last inequality follows from Lemma B.6 and (B.41) in Condition 1. Note that $\left\|\boldsymbol{W}_{:b}^T \boldsymbol{X}\boldsymbol{V}\right\| = \|\boldsymbol{S}_{b:}\| \geq c_3$ and $\|\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X}\boldsymbol{V}\|$ $\geq C_1$ by inequality (B.89). Therefore, we have

$$\begin{aligned}
I_{12} &\lesssim \|(\boldsymbol{W}_{:b}^T - \boldsymbol{W}_{:b}^{(t),T})\boldsymbol{E}\boldsymbol{V}\| \\
&\qquad + \|(\boldsymbol{W}_{:b}^T - \boldsymbol{W}_{:b}^{(t),T})\boldsymbol{X}\boldsymbol{V}\| \|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{E}\boldsymbol{V}\| \\
&\lesssim \sqrt{\frac{r^{2K} + pr^{K+1}}{p^K}} \frac{\sqrt{L^{(t)}}}{\Delta_{\min}} + \frac{\sqrt{rL^{(t)}}}{\Delta_{\min}} \sqrt{\frac{r^{2K}}{p^K}} \\
&\lesssim \sqrt{\frac{r^{2K} + pr^{K+1}}{p^K}} \frac{\sqrt{L^{(t)}}}{\Delta_{\min}},
\end{aligned} \tag{B.93}$$

where second inequality follows from the inequalities (B.91), (B.92), and (B.37) in Condition 1.

Hence, combining inequalities (B.90) and (B.93) yields

$$I_1 \lesssim \frac{rL^{(t)}}{\Delta_{\min}} + \sqrt{\frac{r^{2K} + pr^{K+1}}{p^K}} \frac{\sqrt{L^{(t)}}}{\Delta_{\min}}. \tag{B.94}$$

For I_2 and I_3, recall that $\left\|\boldsymbol{W}_{:b}^T \boldsymbol{X}\boldsymbol{V}\right\| = \|\boldsymbol{S}_{b:}\| \geq c_3$ and $\|\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X}\boldsymbol{V}\| \geq C_1$

by inequality (B.89). By triangle inequality and (B.37) in Condition 1, we have

$$I_2 \leq \frac{\|\boldsymbol{W}_{:b}^T \boldsymbol{E}\boldsymbol{V}\|}{\|\boldsymbol{W}_{:b}^T \boldsymbol{X}\boldsymbol{V}\|} \lesssim \|\boldsymbol{W}_{:b}^T \boldsymbol{E}\boldsymbol{V}\| \lesssim \frac{r^K}{p^{K/2}}, \tag{B.95}$$

and

$$I_3 \leq \frac{\|\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{E}\boldsymbol{V}\|}{\|\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X}\boldsymbol{V}\|} \lesssim \|\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{E}\boldsymbol{V}\| \lesssim \frac{r^K}{p^{K/2}}. \tag{B.96}$$

Therefore, combining the inequalities (B.94), (B.95), and (B.96), we finish the proof of inequality (B.83).

wivde Inequality (B.84). Here we only show the proof of inequality (B.84) with $\boldsymbol{W}_{:b}^{(t)}$. The proof also holds by replacing $\boldsymbol{W}_{:b}^{(t)}$ to $\boldsymbol{W}_{:b}$, and we omit the repeated procedures.

We upper bound the desired quantity by triangle inequality

$$\|[\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{Y}\boldsymbol{V}]^s - [\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\| \leq J_1 + J_2 + J_3,$$

where

$$\begin{aligned}
J_1 &= \left\| \frac{\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{Y}\boldsymbol{V}}{\|\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X}\boldsymbol{V}\|} - \frac{\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{Y}\boldsymbol{V}^{(t)}}{\|\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X}\boldsymbol{V}^{(t)}\|} \right\|, \\
J_2 &= \left\| \left(\frac{1}{\|\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{Y}\boldsymbol{V}\|} - \frac{1}{\|\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X}\boldsymbol{V}\|} \right) \boldsymbol{W}_{:b}^{(t),T} \boldsymbol{Y}\boldsymbol{V} \right\|, \\
J_3 &= \left\| \left(\frac{1}{\|\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{Y}\boldsymbol{V}^{(t)}\|} - \frac{1}{\|\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X}\boldsymbol{V}^{(t)}\|} \right) \boldsymbol{W}_{:b}^{(t),T} \boldsymbol{Y}\boldsymbol{V}^{(t)} \right\|.
\end{aligned}$$

Next, we upper bound the quantities J_1, J_2, J_3 separately.

For J_1, by triangle inequality, we have

$$J_1 \leq J_{11} + J_{12},$$

where

$$J_{11} = \left\| \frac{\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X} \boldsymbol{V}}{\|\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X} \boldsymbol{V}\|} - \frac{\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X} \boldsymbol{V}^{(t)}}{\|\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X} \boldsymbol{V}^{(t)}\|} \right\|$$

and

$$J_{12} = \left\| \frac{\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{E} \boldsymbol{V}}{\|\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X} \boldsymbol{V}\|} - \frac{\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{E} \boldsymbol{V}^{(t)}}{\|\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X} \boldsymbol{V}^{(t)}\|} \right\|.$$

We first consider J_{11}. Define the matrix $\boldsymbol{V}^k := \boldsymbol{W}^{\otimes(k-1)} \otimes \boldsymbol{W}^{(t),\otimes(K-k)}$ for $k = 2, \ldots, K-1$, and denote $\boldsymbol{V}^1 = \boldsymbol{V}^{(t)}$, $\boldsymbol{V}^K = \boldsymbol{V}$. Also, define the quantity

$$J_{11}^k = \|[\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X} \boldsymbol{V}^k]^s - [\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X} \boldsymbol{V}^{k+1}]^s\|,$$

for $k = 1, \ldots, K-1$. Let β_k denote the angle between $\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X} \boldsymbol{V}^k$ and $\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X} \boldsymbol{V}^{k+1}$. With the same idea to prove I_{11} in inequality (B.90), we bound J_{11}^k by the trigonometric function of β_k.

To roughly estimate the range of β_k, we consider the inner product between $\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X} \boldsymbol{V}^k$ and $\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X} \boldsymbol{V}^{k+1}$. Before the specific derivation of the inner product, note that

$$\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X} \boldsymbol{V}^k = \text{Mat}_1(\mathcal{T}_k), \quad \boldsymbol{W}_{:b}^{(t),T} \boldsymbol{X} \boldsymbol{V}^{k+1} = \text{Mat}_1(\mathcal{T}_{k+1}),$$

where

$$\begin{aligned} \mathcal{T}_k = \mathcal{X} \times_1 \boldsymbol{W}_{:b}^{(t),T} \times_2 \boldsymbol{W}^T \times_3 \cdots \times_k \boldsymbol{W}^T \\ \times_{k+1} \boldsymbol{W}^{(t),T} \times_{k+2} \cdots \times_K \boldsymbol{W}^{(t),T} \\ \mathcal{T}_{k+1} = \mathcal{X} \times_1 \boldsymbol{W}_{:b}^{(t),T} \times_2 \boldsymbol{W}^T \times_3 \cdots \times_k \boldsymbol{W}^T \\ \times_{k+1} \boldsymbol{W}^T \times_{k+2} \cdots \times_K \boldsymbol{W}^{(t),T}. \end{aligned}$$

Recall the definition of confusion matrix $\boldsymbol{D} = \boldsymbol{M}^T \boldsymbol{\Theta}^T \boldsymbol{W}^{(t)} = [\![D_{ab}]\!] \in \mathbb{R}^{r \times r}$. We have

$$
\begin{aligned}
&\left\langle \boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}^{k}, \boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}^{k+1}\right\rangle \\
&= \langle \mathrm{Mat}_{k+1}(\mathcal{T}_k), \mathrm{Mat}_{k+1}(\mathcal{T}_{k+1})\rangle \\
&= \left\langle \boldsymbol{D}^{T}\boldsymbol{S}\boldsymbol{Z}^{k}, \boldsymbol{S}\boldsymbol{Z}^{k}\right\rangle \\
&= \sum_{b\in[r]}\left(D_{bb}\|\boldsymbol{S}_{b:}\boldsymbol{Z}^{k}\|^2 + \sum_{a\neq b, a\in[r]} D_{ab}\left\langle \boldsymbol{S}_{a:}\boldsymbol{Z}^{k}, \boldsymbol{S}_{b:}\boldsymbol{Z}^{k}\right\rangle\right) \\
&\gtrsim (1-\log p^{-1})\min_{a\in[r]}\|\boldsymbol{S}_{a:}\boldsymbol{Z}^{k}\|^2 - \log p^{-1}\max_{a\in[r]}\|\boldsymbol{S}_{a:}\boldsymbol{Z}^{k}\|^2, \qquad \text{(B.97)}
\end{aligned}
$$

where $\boldsymbol{Z}^{k} = \boldsymbol{D}_{:b}\otimes \boldsymbol{I}_r^{\otimes(k-1)}\otimes \boldsymbol{D}^{\otimes(K-k-1)}$, the equations follow by the tensor algebra and definitions, and the last inequality follows from the bounds of $\boldsymbol{D}$ in (B.86) and (B.87).

Note that

$$
\begin{aligned}
\|\boldsymbol{D}\|_{\sigma} &\leq \|\boldsymbol{D}\|_{F} \\
&\leq \sqrt{\sum_{b\in[r]} D_{bb}^2 + (\sum_{a\neq b, a,b\in[r]} D_{ab})^2} \\
&\lesssim \sqrt{r+\log^2 p^{-1}} \lesssim 1, \qquad \text{(B.98)}
\end{aligned}
$$

where the second inequality follows from inequality (B.86), and the fact that for all $b\in[r]$,

$$
D_{bb} \lesssim \frac{r}{p}\sum_{i:\ z(i)=b}\theta(i) \lesssim 1.
$$

Also, we have

$$
\lambda_r(\boldsymbol{D}) \geq \lambda_r(\boldsymbol{W}^{(t)})\lambda_r(\boldsymbol{\Theta}\boldsymbol{M}) \gtrsim 1, \qquad \text{(B.99)}
$$

following the Lemma B.6 and Lemma B.10. Then, for all $k\in[K]$, we have

$$
\begin{aligned}
1 &\lesssim \|\boldsymbol{D}_{:b}\|\ \lambda_r(\boldsymbol{D})^{K-k-1} \leq \lambda_{r^{K-2}}(\boldsymbol{Z}^{k}) \\
&\leq \|\boldsymbol{Z}^{k}\|_{\sigma} \leq \|\boldsymbol{D}_{:b}\|\ \|\boldsymbol{D}\|_{\sigma}^{K-k-1} \lesssim 1. \qquad \text{(B.100)}
\end{aligned}
$$

Thus, we have bounds

$$\max_{a\in[r]}\|\boldsymbol{S}_{a:}\boldsymbol{Z}^k\| \le \max_{a\in[r]}\|\boldsymbol{S}_{a:}\|\,\|\boldsymbol{Z}^k\|_\sigma \lesssim 1,$$

$$\min_{a\in[r]}\|\boldsymbol{S}_{a:}\boldsymbol{Z}^k\| \ge \min_{a\in[r]}\|\boldsymbol{S}_{a:}\|\,\lambda_{r^{K-2}}(\boldsymbol{Z}^k) \gtrsim 1.$$

Hence, when p is large enough, the inner product (B.97) is positive, which implies $\beta_k \in [0, \pi/2)$ and thus $2\sin\frac{\beta_k}{2} \le \sqrt{2}\sin\beta_k$.

Next, we upper bound the trigonometric function $\sin\beta_k$. Note that

$$\begin{aligned}\sin\beta_k &= \sin(\boldsymbol{D}_{:b}^T\boldsymbol{S}\boldsymbol{I}_r^{\otimes k-1}\otimes\boldsymbol{D}^{\otimes K-k}, \boldsymbol{D}_{:b}^T\boldsymbol{S}\boldsymbol{I}_r^{\otimes k}\otimes\boldsymbol{D}^{\otimes K-k-1})\\ &\le \sin\beta_{k1} + \sin\beta_{k2},\end{aligned}$$

where

$$\begin{aligned}\sin\beta_{k1} = \sin(&\boldsymbol{D}_{:b}^T\boldsymbol{S}\boldsymbol{I}_r^{\otimes k-1}\otimes\boldsymbol{D}^{\otimes K-k},\\ &\boldsymbol{D}_{:b}^T\boldsymbol{S}\boldsymbol{I}_r^{\otimes k-1}\otimes\tilde{\boldsymbol{D}}\otimes\boldsymbol{D}^{\otimes K-k-1}),\\ \sin\beta_{k2} = \sin(&\boldsymbol{D}_{:b}^T\boldsymbol{S}\boldsymbol{I}_r^{\otimes k-1}\otimes\tilde{\boldsymbol{D}}\otimes\boldsymbol{D}^{\otimes K-k-1},\\ &\boldsymbol{D}_{:b}^T\boldsymbol{S}\boldsymbol{I}_r^{\otimes k}\otimes\boldsymbol{D}^{\otimes K-k-1}),\end{aligned}$$

and $\tilde{\boldsymbol{D}}$ is the normalized confusion matrix with entries $\tilde{\boldsymbol{D}}_{ab} = \frac{\sum_{i\in[p]}\theta(i)\mathbb{1}\{z^{(t)}=b, z(i)=a\}}{\sum_{i\in[p]}\theta(i)\mathbb{1}\{z^{(t)}=b\}}$.

To bound $\sin\beta_{k1}$, recall Definition 3.16 that for any cluster assignment $\bar{z}$ in the ε-neighborhood of true z,

$$\begin{aligned}\boldsymbol{p}(\bar{z}) &= (|\bar{z}^{-1}(1)|, \ldots, |\bar{z}^{-1}(r)|)^T,\\ \boldsymbol{p}_{\boldsymbol{\theta}}(\bar{z}) &= (\|\boldsymbol{\theta}_{\bar{z}^{-1}(1)}\|_1, \ldots, \|\boldsymbol{\theta}_{\bar{z}^{-1}(r)}\|_1)^T.\end{aligned}$$

Note that we have $\ell^{(t)} \le \frac{L^{(t)}}{\Delta_{\min}^2} \le \frac{\bar{C}}{\bar{C}} r\log^{-1}(p)$ by Condition 1 and Lemma 3.11. Then, with the locally linear stability assumption, the $\boldsymbol{\theta}$ is $\ell^{(t)}$-locally linearly

stable; i.e.,

$$\sin(\boldsymbol{p}(z^{(t)}), \boldsymbol{p}_{\boldsymbol{\theta}}(z^{(t)})) \lesssim \frac{L^{(t)}}{\Delta_{\min}}.$$

Note that $\operatorname{diag}(\boldsymbol{p}(z^{(t)}))\boldsymbol{D} = \operatorname{diag}(\boldsymbol{p}_{\boldsymbol{\theta}}(z^{(t)}))\tilde{\boldsymbol{D}}$, and $\sin(\boldsymbol{a}, \boldsymbol{b}) = \min_{c\in\mathbb{R}} \frac{\|\boldsymbol{a}-c\boldsymbol{b}\|}{\|\boldsymbol{a}\|}$ for vectors $\boldsymbol{a}, \boldsymbol{b}$ of same dimension. Let $c_0 = \arg\min_{c\in\mathbb{R}} \frac{\|\boldsymbol{p}(z^{(t)})-c\boldsymbol{p}_{\boldsymbol{\theta}}(z^{(t)})\|}{\|\boldsymbol{p}(z^{(t)})\|}$. Then, we have

$$\begin{aligned}
&\min_{c\in\mathbb{R}} \|\boldsymbol{D} - c\tilde{\boldsymbol{D}}\|_F \\
&\leq \|\boldsymbol{I}_r - c_0 \operatorname{diag}(\boldsymbol{p}(z^{(t)}))\operatorname{diag}^{-1}(\boldsymbol{p}_{\boldsymbol{\theta}}(z^{(t)}))\|_F \|\boldsymbol{D}\|_F \\
&\lesssim \frac{\|\boldsymbol{p}(z^{(t)}) - c_0 \boldsymbol{p}_{\boldsymbol{\theta}}(z^{(t)})\|}{\min_{a\in[r]} \|\boldsymbol{\theta}_{z^{(t),-1}(a)}\|_1} \\
&= \frac{\|\boldsymbol{p}(z^{(t)})\|}{\min_{a\in[r]} \|\boldsymbol{\theta}_{z^{(t),-1}(a)}\|_1} \sin(\boldsymbol{p}(z^{(t)}), \boldsymbol{p}_{\boldsymbol{\theta}}(z^{(t)})) \\
&\lesssim \frac{L^{(t)}}{\Delta_{\min}},
\end{aligned}$$

where the last inequality follows from Lemma B.10, the constraint $\min_{i\in[p]} \theta(i) \geq c > 0$, $\|\boldsymbol{p}(z^{(t)})\| \lesssim p$ and $\min_{a\in[r]} \|\boldsymbol{\theta}_{z^{(t),-1}(a)}\|_1 \gtrsim p$.

By the geometry property of trigonometric function, we have

$$\begin{aligned}
\sin\beta_{k1} &= \min_{c\in\mathbb{R}} \frac{\|\boldsymbol{D}_{:b}^T \boldsymbol{S}\boldsymbol{I}_r^{\otimes k-1} \otimes (\boldsymbol{D} - c\tilde{\boldsymbol{D}}) \otimes \boldsymbol{D}^{\otimes K-k-1}\|}{\|\boldsymbol{D}_{:b}^T \boldsymbol{S}\boldsymbol{I}_r^{\otimes k-1} \otimes \boldsymbol{D}^{\otimes K-k}\|} \\
&\leq \frac{\|\boldsymbol{D}_{:b}^T \boldsymbol{S}\| \|\boldsymbol{D} - c_0\tilde{\boldsymbol{D}}\|_\sigma \|\boldsymbol{D}\|_\sigma^{K-k-1}}{\|\boldsymbol{D}_{:b}^T \boldsymbol{S}\| \lambda_r^{K-k}(\boldsymbol{D})} \\
&\lesssim \|\boldsymbol{D} - c_0\tilde{\boldsymbol{D}}\|_F \\
&\lesssim \frac{L^{(t)}}{\Delta_{\min}},
\end{aligned} \tag{B.101}$$

where the second inequality follows from the singular property of $\boldsymbol{D}$ in (B.98), (B.99) and the constraint of $\boldsymbol{S}$ in (3.2).

To bound $\sin\beta_{k2}$, let $\boldsymbol{C} = \text{diag}(\{\|\boldsymbol{S}_{a:}\|\}_{a\in[r]})$. We have

$$\begin{aligned}
\sin\beta_{k2} &\lesssim \frac{\left\|\boldsymbol{D}_{:b}^T\boldsymbol{S}\boldsymbol{I}_r^{\otimes k-1}\otimes(\boldsymbol{I}_r-\tilde{\boldsymbol{D}})\otimes\boldsymbol{D}^{\otimes K-k-1}\right\|}{\|\boldsymbol{D}_{:b}^T\boldsymbol{S}\boldsymbol{I}_r^{\otimes k}\otimes\boldsymbol{D}^{\otimes K-k-1}\|} \\
&\lesssim \frac{\|(\boldsymbol{I}_r-\tilde{\boldsymbol{D}}^T)\boldsymbol{S}\boldsymbol{Z}^k\|_F}{\|\boldsymbol{D}_{:b}^T\boldsymbol{S}\|\lambda_r^{K-k-1}(\boldsymbol{D})} \\
&\lesssim \|(\boldsymbol{I}_r-\tilde{\boldsymbol{D}}^T)\boldsymbol{S}\boldsymbol{C}^{-1}\|_F\|\boldsymbol{C}\boldsymbol{Z}^k\|_\sigma \\
&\lesssim \frac{r}{p}\sum_{i\in[p]}\theta(i)\sum_{b\in[r]}\mathbb{1}\{z^{(t)}(i)=b\}\|\boldsymbol{S}_{b:}^s-\boldsymbol{S}_{z(i):}^s\| \\
&\lesssim \frac{L^{(t)}}{\Delta_{\min}},
\end{aligned} \tag{B.102}$$

where the third inequality follows from the singular property of $\boldsymbol{D}$ and the boundedness of $\boldsymbol{S}$, and the fourth inequality follows from the definition of $\tilde{\boldsymbol{D}}$, boundedness of $\boldsymbol{S}$, the lower bound of $\boldsymbol{\theta}$, and the singular property of $\boldsymbol{Z}^k$ in inequality (B.100), and the last line follows from the definition of $L^{(t)}$.

Combining (B.101) and (B.102) yields

$$\sin\beta_k \le \sin\beta_{k1}+\sin\beta_{k2} \lesssim \frac{L^{(t)}}{\Delta_{\min}}.$$

Finally, by triangle inequality, we obtain

$$J_{11} \le \sum_{k=1}^{K-1} J_{11}^k \lesssim \sum_{k=1}^{K-1}\sin\beta_k \lesssim (K-1)\frac{rL^{(t)}}{\Delta_{\min}}. \tag{B.103}$$

We now consider J_{12}. By triangle inequality, we have

$$\begin{aligned}
J_{12} \le{}& \frac{1}{\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}\|}\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{E}(\boldsymbol{V}-\boldsymbol{V}^{(t)})\| \\
&+\frac{\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}(\boldsymbol{V}-\boldsymbol{V}^{(t)})\|}{\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}\|\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}^{(t)}\|}\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{E}\boldsymbol{V}^{(t)}\|.
\end{aligned}$$

Note that

$$\begin{aligned}\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}^{(t)}\| &= \|\boldsymbol{D}^T\boldsymbol{S}\boldsymbol{Z}^1\| \\ &\geq \lambda_r(\boldsymbol{D})\,\|\boldsymbol{S}\|\,\lambda_{r^{K-2}}(\boldsymbol{Z}^1) \gtrsim 1, \end{aligned} \tag{B.104}$$

where the inequality follows from the bounds (B.99) and (B.100).

By (Han et al., 2022a, Proof of Lemma 5), we have

$$\begin{aligned}&\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{E}(\boldsymbol{V}-\boldsymbol{V}^{(t)})\| \\ &\qquad \lesssim \sqrt{\frac{r^{2K+1}+pr^{2+K}}{p^K}}\frac{(K-1)\sqrt{L^{(t)}}}{\Delta_{\min}}. \end{aligned} \tag{B.105}$$

Notice that

$$\begin{aligned}&\|\boldsymbol{X}(\boldsymbol{V}^k-\boldsymbol{V}^{k+1})\|_F \\ &\leq \|(\boldsymbol{I}-\boldsymbol{D}^T)\boldsymbol{S}(\boldsymbol{I}_r^{\otimes(k-1)}\otimes\boldsymbol{D}^{\otimes(K-k-1)})\|_F \\ &\leq \|(\boldsymbol{W}^T-\boldsymbol{W}^{(t),T})\boldsymbol{\Theta}\boldsymbol{M}\|_F\,\|\boldsymbol{S}\|_F\,\|\boldsymbol{D}\|_\sigma^{K-k-1} \\ &\lesssim \|\boldsymbol{W}^T-\boldsymbol{W}^{(t),T}\|\,\|\boldsymbol{\Theta}\boldsymbol{M}\|_\sigma \\ &\lesssim \frac{\sqrt{rL^{(t)}}}{\Delta_{\min}}, \end{aligned} \tag{B.106}$$

where the first inequality follows from the tensor algebra in inequality (B.97), the second inequality follows from the fact that $\boldsymbol{I}=\boldsymbol{W}^T\boldsymbol{\Theta}\boldsymbol{M}$, and the last inequality follows from (Han et al., 2022a, Proof of Lemma 5). It follows from (B.106) and Lemma B.10 that

$$\begin{aligned}\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}(\boldsymbol{V}-\boldsymbol{V}^{(t)})\| &\leq \|\boldsymbol{W}_{:b}^{(t),T}\|\sum_{k=1}^{K-1}\|\boldsymbol{X}(\boldsymbol{V}^k-\boldsymbol{V}^{k+1})\|_F \\ &\lesssim \frac{\sqrt{rL^{(t)}}}{\sqrt{p}\Delta_{\min}}. \end{aligned} \tag{B.107}$$

Note that $\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}\|$ and $\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}^{(t)}\|$ are lower bounded by inequal-

ities (B.89) and (B.104), respectively. We have

$$\begin{aligned} J_{12} &\lesssim \|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{E}(\boldsymbol{V}-\boldsymbol{V}^{(t)})\| \\ &\qquad + \|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}(\boldsymbol{V}-\boldsymbol{V}^{(t)})\|\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{E}\boldsymbol{V}^{(t)}\| \\ &\lesssim \sqrt{\frac{r^{2K+1}+pr^{2+K}}{p^K}}\frac{\sqrt{L^{(t)}}}{\Delta_{\min}} + \frac{\sqrt{rL^{(t)}}}{\sqrt{p}\Delta_{\min}}\sqrt{\frac{r^{2K}}{p^K}} \\ &\lesssim \sqrt{\frac{r^{2K+1}+pr^{2+K}}{p^K}}\frac{\sqrt{L^{(t)}}}{\Delta_{\min}}, \end{aligned}$$

where the second inequality follows from inequalities (B.105), (B.107), and the inequality (B.37) in Condition 1.

For J_2 and J_3, recall that $\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}\|$ and $\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}^{(t)}\|$ are lower bounded by inequalities (B.89) and (B.104), respectively. By triangle inequality and inequality (B.37) in Condition 1, we have

$$J_2 \le \frac{\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{E}\boldsymbol{V}\|}{\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}\|} \lesssim \|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{E}\boldsymbol{V}\| \lesssim \frac{r^K}{p^{K/2}}, \tag{B.108}$$

and

$$J_3 \le \frac{\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{E}\boldsymbol{V}^{(t)}\|}{\|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{X}\boldsymbol{V}^{(t)}\|} \lesssim \|\boldsymbol{W}_{:b}^{(t),T}\boldsymbol{E}\boldsymbol{V}\| \lesssim \frac{r^K}{p^{K/2}}. \tag{B.109}$$

Therefore, combining the inequalities (B.103), (B.108), and (B.109), we finish the proof of inequality (B.84).

wvde Inequality (B.85). By triangle inequality, we upper bound the desired

quantity

$$\begin{aligned}
&\|[\boldsymbol{W}_{:b}^T \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\| \\
&\le \|[\boldsymbol{W}_{:b}^T \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s - [\boldsymbol{W}_{:b}^T \boldsymbol{Y}\boldsymbol{V}]^s\| \\
&\qquad + \|[\boldsymbol{W}_{:b}^T \boldsymbol{Y}\boldsymbol{V}]^s - [\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{Y}\boldsymbol{V}]^s\| \\
&\qquad + \|[\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{Y}\boldsymbol{V}]^s - [\boldsymbol{W}_{:b}^{(t),T} \boldsymbol{Y}\boldsymbol{V}^{(t)}]^s\| \\
&\lesssim \frac{rL^{(t)}}{\Delta_{\min}} + \sqrt{\frac{rr^{2K} + pr^{K+2}}{p^K}} \frac{\sqrt{L^{(t)}}}{\Delta_{\min}},
\end{aligned}$$

following the inequalities (B.83) and (B.84). Therefore, we finish the proof of inequality (B.85).

Next, we show the intermediate inequalities holds with $\boldsymbol{P}, \boldsymbol{Q}$ and $L(\hat{z})$. Consider the MLE confusion matrix $\hat{\boldsymbol{D}} = \boldsymbol{M}^T \boldsymbol{\Theta}^T \hat{\boldsymbol{P}} = [\![\hat{D}_{ab}]\!] \in \mathbb{R}^{r\times r}$ with entries

$$\begin{aligned}
\hat{D}_{ab} &= \frac{\sum_{i\in[p]} \theta(i)\hat{\theta}(i) \mathbb{1}\{z(i) = a, \hat{z}(i) = b\}}{\|\hat{\boldsymbol{\theta}}_{\hat{z}^{-1}(b)}\|^2} \\
&= \frac{\sum_{i\in[p]} (1 + o(p^{K-2}))(\hat{\theta}(i))^2 \mathbb{1}\{z(i) = a, \hat{z}(i) = b\}}{\|\hat{\boldsymbol{\theta}}_{\hat{z}^{-1}(b)}\|^2},
\end{aligned} \tag{B.110}$$

where the second equation follows from Lemma B.13, and thus $\sum_{a\in[r]} \hat{\boldsymbol{D}}_{ab} = 1 + o(1)$. By the derivation of (B.86), (B.87), (B.99), and (B.98), we have

$$\sum_{a\neq b\in[r]} \hat{D}_{ab} \lesssim \frac{1}{p} \sum_{i\in[p]} \mathbb{1}\{\hat{z}(i) \neq z(i)\}(\hat{\theta}(i))^2 \lesssim \frac{1}{\log p},$$

$$\hat{D}_{bb} \gtrsim 1 - \frac{1}{\log p}, \quad \lambda_{\min}(\hat{\boldsymbol{D}}) \asymp \|\hat{\boldsymbol{D}}\|_\sigma = (1 + o(1)).$$

for all $a \neq b \in [r]$.

Now, we are ready to show the intermediate inequalities. First, by Lemma 3.4 and $\min_{i\in[p]} \theta(i) \geq c$, we have

$$\|\boldsymbol{S}_{a:}^s - \boldsymbol{S}_{b:}^s\| \asymp \|\boldsymbol{A}_{a:}^s - \boldsymbol{A}_{b:}^s\|.$$

Then we can replace the $L^{(t)}$ by $L(\hat{z})$ in the proof of Lemma B.12. The analogies of inequalities (B.81), (B.82), (B.83), (B.84), and (B.85) hold by using the MLE confusion matrix and the definition of $L(\hat{z})$.

Particularly, for the analogy of (B.84), the usage of MLE confusion matrix avoids the stability condition on $\boldsymbol{\theta}$. Let $\bar{\boldsymbol{D}}$ be the normalized version of $\hat{\boldsymbol{D}}$. The angle in inequality (B.101) decays to 0 at speed $p^{-(K-2)} \lesssim \Delta_{\min}$ when $K \geq 3$, and the inequality (B.102) holds by the fact that

$$\begin{aligned}\|(\boldsymbol{I}_r - \bar{\boldsymbol{D}})\boldsymbol{S}\boldsymbol{C}^{-1}\|_F &\lesssim \frac{r}{p}\sum_{i\in[p]}(\theta(i))^2 \sum_{b\in[r]}\|\boldsymbol{S}^s_{b:} - \boldsymbol{S}^s_{z(i):}\| \\ &\lesssim \frac{r}{p}\sum_{i\in[p]}(\theta(i))^2 \sum_{b\in[r]}\|\boldsymbol{A}^s_{b:} - \boldsymbol{A}^s_{z(i):}\|.\end{aligned}$$

□

Lemma B.13 (Polynomial estimation error of MLE)**.** *Let $(\hat{z}, \hat{\mathcal{S}}, \hat{\boldsymbol{\theta}})$ denote the MLE in* (3.9) *with fixed $K \geq 2$ and symmetric mean tensor, and $\hat{\mathcal{X}}$ denote the mean tensor consisting of parameter $(\hat{z}, \hat{\mathcal{S}}, \hat{\boldsymbol{\theta}})$. With high probability going to 1 as $p \to \infty$, we have*

$$\|\mathcal{X} - \hat{\mathcal{X}}\|_F^2 \lesssim \sigma^2\left(r^K + Kpr\right),$$

with probability going to 1. When SNR $\gtrsim p^{-(K-1)}\log p$, $\boldsymbol{\theta}$ is balanced, and $\min_{i\in[p]}\theta(i) \geq c$ for some positive constant c, the MLE satisfies

$$\begin{aligned}&\frac{1}{p}\sum_{i\in[p]}\mathbb{1}\{\hat{z}(i) \neq z(i)\}(\theta(i))^2 \lesssim \frac{1}{r\log p}, \\ &\frac{1}{p}\sum_{i\in[p]}\mathbb{1}\{\hat{z}(i) \neq z(i)\}(\hat{\theta}(i))^2 \lesssim \frac{1}{r\log p}, \\ &\text{and } L(\hat{z}) \lesssim \frac{\Delta_{\min}^2}{r\log p},\end{aligned}$$

Further, we have

$$\theta(i)^2 = (1 + o(p^{-(K-2)}))\hat{\theta}(i)^2.$$

Proof of Lemma B.13. Without loss of generality, we assume $\sigma^2 = 1$ and identity

mapping minimizes the misclustering error for MLE. For arbitrary two sets of parameters $(z, \mathcal{S}, \boldsymbol{\theta}), (z', \mathcal{S}', \boldsymbol{\theta}') \in \mathcal{P}(\gamma)$ and corresponding mean tensors $\mathcal{X}, \mathcal{X}'$, we have

$$\begin{aligned}
&\mathsf{rank}(\mathsf{Mat}_k(\mathcal{X}) - \mathsf{Mat}_k(\mathcal{X}')) \\
&\quad \leq \mathsf{rank}(\mathsf{Mat}_k(\mathcal{X})) + \mathsf{rank}(\mathsf{Mat}_k(\mathcal{X})) \\
&\quad \leq 2r, \quad k \in [K].
\end{aligned}$$

Hence, we have

$$\mathcal{X} - \mathcal{X}' \in \mathcal{Q}(2r, \ldots, 2r), \tag{B.111}$$

where $\mathcal{Q}(r, \ldots, r) \coloneqq \{\text{Tucker tensor with rank } (r, \ldots, r)\}$.

Then, we obtain that

$$\begin{aligned}
&\mathbb{P}(\|\mathcal{X} - \hat{\mathcal{X}}_{ML}\|_F \geq t) \\
&\leq 2\mathbb{P}\left(\sup_{\mathcal{X}, \mathcal{X}' \in \mathcal{P}(r,\ldots,r)} \left\langle \frac{\mathcal{X} - \mathcal{X}'}{\|\mathcal{X} - \mathcal{X}'\|_F}, \mathcal{E} \right\rangle \geq t\right) \\
&\leq 2\mathbb{P}\left(\sup_{\mathcal{T} \in \mathcal{Q}(2r,\ldots,2r) \cap \{\|\mathcal{T}\|_F = 1\}} \langle \mathcal{T}, \mathcal{E} \rangle \geq t\right) \\
&\lesssim \exp(-Kpr),
\end{aligned}$$

with the choice $t \asymp \sigma\sqrt{(Kpr + r^K)}$. Here the first inequality follows from (Wang and Zeng, 2019, Lemma 1), the second inequality follows from (B.111), and the last inequality follows from (Han et al., 2022b, Lemma E5).

When $\Delta_{\min}^2 \gtrsim p^{-(K-1)} \log p$, we replace the vector $\hat{x}_{\hat{z}(i)}$ and $\hat{\boldsymbol{X}}$ by our MLE estimator in the proof of Theorem 3.12. With estimation error $\|\mathcal{X} - \hat{\mathcal{X}}\|_F^2 \lesssim \left(r^K + Kpr\right)$

and $\Delta_{\min}^2 \gtrsim p^{-(K-1)} \log p$, we have

$$\begin{aligned}\frac{1}{p}\sum_{i\in[p]} \mathbb{1}\{\hat{z}(i) \neq z(i)\}(\theta(i))^2 &\lesssim \frac{r^{K-1}}{\Delta_{\min}^2 p^K}\|\mathcal{X} - \hat{\mathcal{X}}\|_F^2 \\ &\lesssim \frac{r^{K-2}}{p^{K-1}\Delta_{\min}^2} \\ &\lesssim \frac{1}{r\log p},\end{aligned}$$

and

$$L(\hat{z}) \lesssim \frac{\Delta_{\min}^2}{r\log p}.$$

Above result holds for $\hat{\theta}(i)$ after switching the parameters $\boldsymbol{X}$ with $\hat{\boldsymbol{X}}$ and switch $\boldsymbol{\theta}$ with $\hat{\boldsymbol{\theta}}$ in the proof.

Last, notice that for all $a \in [r]$

$$\begin{aligned}&(1-O(1))\frac{p^2}{r^2}\|\boldsymbol{W}_{:a}^T\boldsymbol{X} - \hat{\boldsymbol{W}}_{:a}^T\hat{\boldsymbol{X}}\|_F^2 \\ &\leq \|\sum_{\hat{z}(i)=z(i)=a}(\theta(i)\boldsymbol{W}_{:a}^T\boldsymbol{X} - \hat{\theta}(i)\hat{\boldsymbol{W}}_{:a}^T\hat{\boldsymbol{X}})\|_F^2 \\ &\leq \|\mathcal{X} - \hat{\mathcal{X}}\|_F^2 \leq pr,\end{aligned}$$

where the first inequality follows from the facts that $\ell(\hat{z}, z) \lesssim \frac{1}{\log p}$, $|z^{-1}(a)| \asymp p/r$,

$$|z^{-1}(a)| - C\frac{p}{r}\ell(\hat{z}, z) \leq |\hat{z}^{-1}(a)| \leq |z^{-1}(a)| + C\frac{p}{r}\ell(\hat{z}, z),$$

$$|z^{-1}(a)| - C\frac{p}{r}\ell(\hat{z}, z) \leq \sum_{z(i)=z(i)=a}\theta(i) \leq |z^{-1}(a)|,$$

and

$$|\hat{z}^{-1}(a)| - C\frac{p}{r}\ell(\hat{z}, z) \leq \sum_{\hat{z}(i)=z(i)=a}\hat{\theta}(i) \leq |\hat{z}^{-1}(a)|.$$

Hence, for all $i \in [p]$

$$\begin{aligned}
&(\theta(i) - \hat{\theta}(i))^2 \|\boldsymbol{W}_{:a}^T \boldsymbol{X}\|_F^2 - O(p) \\
&\leq \|(\theta(i) - \hat{\theta}(i))\boldsymbol{W}_{:a}^T \boldsymbol{X}\|_F^2 - \|\hat{\theta}(i)(\boldsymbol{W}_{:a}^T \boldsymbol{X} - \hat{\boldsymbol{W}}_{:a}^T \hat{\boldsymbol{X}})\|_F^2 \\
&\leq \|\mathcal{X} - \hat{\mathcal{X}}\|_F^2 \leq pr,
\end{aligned}$$

where the first inequality follows from $\|\boldsymbol{W}_{:a}^T \boldsymbol{X} - \hat{\boldsymbol{W}}_{:a}^T \hat{\boldsymbol{X}}\|_F^2 \lesssim 1/p$ and $\hat{\theta}(i) \lesssim \frac{p}{r}$. Notice that for all $a \in [r]$

$$\|\boldsymbol{W}_{:a}^T \boldsymbol{X}\|_F^2 \geq \|\boldsymbol{S}_{a:}\|_F^2 \lambda_{\min}^{2(K-1)}(\boldsymbol{\Theta}\boldsymbol{M}) \gtrsim p^{K-1}.$$

The inequality indicates that $\theta(i)^2 = (1 + o(p^{-(K-2)}))\hat{\theta}(i)^2$.

□

C APPENDIX OF CHAPTER 4

C.1 Full TBM Algorithms in Han et al. (2022a)

We present the full TBM algorithm in Han et al. (2022a) for consistency.

Algorithm 5 Full TBM algorithm with spectral initialization and Lloyd iteration Han et al. (2022a)

Input: Observation $\mathcal{Y}$, number of communities $\{r_k\}$, relaxation factor $M > 1$, iteration number T

Higher-order spectral initialization

Compute $\tilde{\boldsymbol{U}}_k = \mathrm{SVD}_{r_k}(\mathrm{Mat}_k(\mathcal{Y}))$ for $k \in [K]$.
for $k = 1$ to K **do**
  Compute the singular space estimator $\hat{\boldsymbol{U}}_k$ via

$$\hat{\boldsymbol{U}}_k = \mathrm{SVD}_{\min\{r_k, r_{-k}\}}\left(\mathrm{Mat}_k(\mathcal{Y} \times_1 \tilde{\boldsymbol{U}}_1^T \times_2 \cdots \times_{k-1} \tilde{\boldsymbol{U}}_{k-1}^T \times_{k+1} \tilde{\boldsymbol{U}}_{k+1}^T \times_{k+2} \cdots \times_K \tilde{\boldsymbol{U}}_k^T)\right)$$

end for
for $k = 1$ to K **do**
  Calculate $\hat{\boldsymbol{Y}}_k = \hat{\boldsymbol{U}}_k \hat{\boldsymbol{U}}_k^T \mathrm{Mat}_k\left(\mathcal{Y} \times_1 \hat{\boldsymbol{U}}_1^T \times_2 \cdots \times_{k-1} \hat{\boldsymbol{U}}_{k-1}^T \times_{k+1} \hat{\boldsymbol{U}}_{k+1}^T \times_{k+2} \cdots \times_K \hat{\boldsymbol{U}}_k^T\right)$
  Find $z_k^{(0)} : [n_k] \mapsto [r_k]$ and centroids $\hat{\boldsymbol{x}}_1, \ldots, \hat{\boldsymbol{x}}_{r_k} \in \mathbb{R}^{r_{-k}}$ such that

$$\sum_{j \in [n_k]} \|\hat{\boldsymbol{Y}}_k(j,:)^T - \hat{\boldsymbol{x}}_{z_k^{(0)}(j)}\|_2^2 \le M \min_{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{r_k} \in \mathbb{R}^{r_{-k}}, z_k : [n_k] \mapsto [r_k]} \sum_{j \in [n_k]} \|\hat{\boldsymbol{Y}}_k(j,:)^T - \boldsymbol{x}_{z_k(j)}\|_2^2.$$

end for
Obtain the spectral initialization $\{z_k^{(0)}\}$.

Higher-order Lloyd algorithm

for $t = 0$ to $T - 1$ **do**
  Update the block means $\mathcal{S}^{(t)}$ via

$$\mathcal{S}^{(t)}(i_1, \ldots, i_K) = \mathrm{Average}\left(\left\{\mathcal{Y}(j_1, \ldots, j_K) : z_k^{(t)}(j_k) = i_k, k \in [K]\right\}\right).$$

  for $k = 1$ to K **do**
    for $j = 1$ to n_k **do**
      Calculate $\mathcal{Y}_k^{(t)} \in \mathbb{R}^{r_1 \times \cdots \times r_{k-1} \times r_{k+1} \times \cdots \times r_K}$ such that

$$\mathcal{Y}_k^{(t)}(i_1, \ldots, i_{k-1}, j, i_{k+1}, \ldots, i_K) = \mathrm{Average}\left(\left\{\mathcal{Y}(j_1, \ldots, j_{k-1}, j, j_{k+1}, \ldots, j_K) : z_l^{(t)}(j_l) = i_l, l \in [K]/k\right\}\right).$$

      Update the mode-k membership for the j-th entry $z_k^{(t+1)}(j)$ via

$$z_k^{(t+1)}(j) = \arg\min_{a \in [r_k]} \|\mathrm{Mat}_k(\mathcal{Y}_k^{(t)})(j,:) - \mathrm{Mat}_k(\mathcal{S}^{(t)})(a,:)\|_2^2.$$

    end for
  end for
end for

Output: Estimated block memberships $\{z_k^{(T)}\}$.

C.2 Additional Numerical Experiments

Additional Simulation

Here, we provide additional simulation results for supplement.

The first additional experiments further investigate the LS-TBM under weak SNR $\mathcal{O}(n^{-3/4})$ cases. Based on Remark 4.8, $m = \mathcal{O}(n^{1/2})$ is the phase transition threshold for exact recovery, and LS-TBM performance is expected to change rapidly around this threshold. We consider two other choices for m, $\mathcal{O}(n^{1/2}\log^{-1/4} n)$ and $\mathcal{O}(n^{1/2}\log^{1/4} n)$, which are slightly smaller and larger than $\mathcal{O}(n^{1/2})$. Figure C.1 shows that the LS-TBM accuracy indeed changes rapidly around the threshold $\mathcal{O}(n^{1/2})$, given that the actual seed sizes associated with the three m values are very close. Moreover, Figure 4.2 demonstrates the divergent error for seeds as the dimension increases. This phenomenon implies that the divergent error of LS-TBM with m equal to or smaller than $\mathcal{O}(n^{1/2})$ is caused by the TBM error on the sub-tensor. Such divergence agrees with the unstable performance of TBM at the phase transition thresholds, as reported in Han et al. (2022a, Figure 3). TBM provides a stable performance only when SNR exceeds the threshold. Nevertheless, above numerical experiments still support the theoretical phase transition in Corollary 4.7, as LS-TBM exhibits sharp changes around the threshold $m = \mathcal{O}(n^{1/2})$.

The second additional experiment explores the empirical LS-TBM performance on tensors of order $K = 3, 4, 5$. Since the number of tensor entries grows exponentially with increasing K, we set $n = 60$ to ensure that the data tensor can be generated within 32 GB of memory. We also consider fixed seed sizes $m = 15, 30, 45$ for LS-TBM due to the small n. Figure C.2 confirms that the computational costs for tensor methods increase exponentially as the tensor order increases. This result aligns with the complexity analysis in Table 4.1. Furthermore, this experiment indicates the importance of our scalable LS-TBM approach for the analysis of higher-order tensors.

Additional results in Uber Pickup data analysis

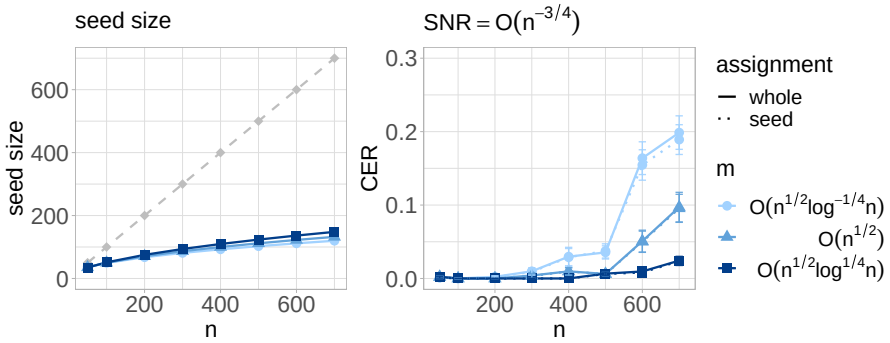

Figure C.1: Additional theoretical verifications of LS-TBM under weak SNR $\mathcal{O}(n^{-3/4})$ cases. In addition to the phase transition threshold $m = \mathcal{O}(n^{1/2})$, we consider two other choices of m for LS-TBM, $\mathcal{O}(n^{1/2}\log^{-1/4} n)$ and $\mathcal{O}(n^{1/2}\log^{1/4} n)$, which are slightly smaller and larger than $\mathcal{O}(n^{1/2})$. (Left) Actual seed sizes associated with different m's. The gray dashed line refers to the dimension n. (Right) CER trajectories as the dimension n from 50 to 700. CER trajectories for seeds (dotted lines) are largely overlapped with those for the LS-TBM Algorithm 4 outputs (solid lines).

Figure C.3 summarizes the Time Cluster results of LS-TBM with $m_1 = 0.1n_1, (m_2, m_3) = 0.5(n_2, n_3)$ and full TBM for the Uber Pickup data. The presented LS-TBM results are obtained from a randomly selected run among multiple executions of the LS-TBM. The CER between the LS-TBM and TBM time assignments is 0.33.

Despite the relatively large CER, LS-TBM captures several key patterns consistent with TBM:

1. Time Cluster 1: The largest cluster with the smallest average Uber pickup amounts, primarily covering early morning hours from midnight to 6 a.m..

2. Time Clusters 2 and 3: Show more even timestamp distributions throughout the daytime.

3. Time Clusters 4, 5 and 6: Smaller clusters capturing evening hours, with Cluster 4 focusing on weekdays, while Clusters 5 and 6 concentrate around

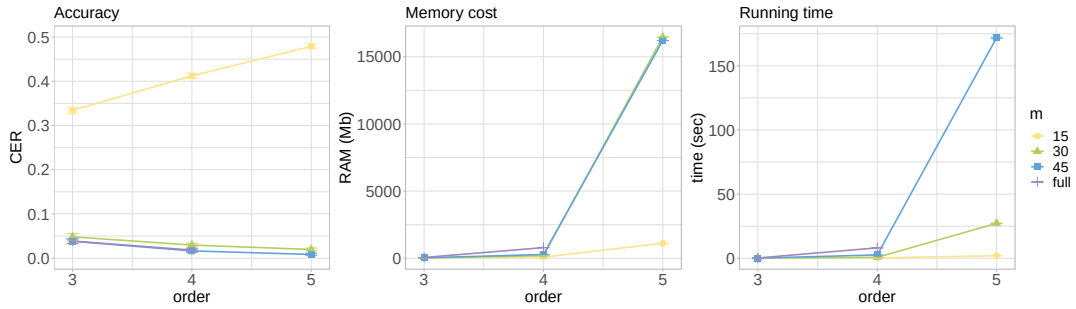

Figure C.2: Empirical comparison between LS-TBMs and full TBM from accuracy (CER), memory cost (peak RAM, in Mb), and running time (in second). We consider dimension $n = 60$, SNR $= \mathcal{O}(1)$, tensor order $K = 3, 4, 5$, and seed size $m = 15, 30, 45$. Full TBM exceeds the memory upper limit of 32 Gb when $K = 5$.

weekends. These clusters exhibit relatively higher pickup amounts than others.

To further evaluate LS-TBM, we consider the identification of Time Cluster 1, labeling other clusters as 0. Using TBM assignments as the ground truth, LS-TBM achieves an accuracy of 0.7, recall of 0.6, and specificity of 0.76. McNemar's test for the confusion matrix yields a p-value of 0.25, indicating no significant difference between LS-TBM and TBM assignments for Cluster 1. This pattern similarity is consistent across multiple runs of LS-TBM. The average Cluster 1 identification accuracy is 0.74 over 10 replications. Hence, we conclude that LS-TBM effectively captures the major clustering patterns of TBM for the Uber Pickup data.

Table C.1 compares the computational performance between LS-TBMs with different m's and TBM in Uber Pickup application. LS-TBM is shown to have a better computation performance than TBM. Additionally, Table C.1 indicates the variability of LS-TBM across runs, with a CER variance of approximately 0.05. Note that TBM keeps the same result across different runs. The LS-TBM variability arises from the inherent randomness in the method, which is a trade-off for its computational speed-up. We leave the extension of LS-TBM incorporating more robust procedures as a future work.

	$a = 0.1$	$a = 0.2$	$a = 0.3$	$a = 0.5$	$a = 0.7$	Full TBM
Averaged CER	0.18 (0.02)	0.15 (0.02)	0.12 (0.05)	0.12 (0.02)	0.08 (0.05)	-
mode-1 CER	0.36 (0.03)	0.34 (0.02)	0.27 (0.09)	0.32 (0.05)	0.19 (0.15)	-
mode-2 CER	0.11 (0.04)	0.06 (0.03)	0.06 (0.07)	0.02 (0.01)	0.03 (0.06)	-
mode-3 CER	0.07 (0.03)	0.05 (0.02)	0.04 (0.04)	0.02 (0.01)	0.01 (0.01)	-
Peak RAM (Mb)	721.26 (13.10)	2858.20 (0.00)	7643.19 (0.03)	11716.60 (0.00)	15869.99 (2116.02)	26853.1
Running time (sec)	1.31 (0.10)	9.05 (0.14)	32.55 (0.18)	188.61 (1.36)	632.74 (1.98)	2335.7

Table C.1: Computational comparison between LS-TBM and TBM in real Uber Pickup tensor application. We set $a = m_k/n_k$ from 0.1 to 0.7 for LS-TBM and take full TBM assignments as reference “true" assignments in CER. Standard deviations across 10 replications for LS-TBM applications are recorded in parentheses.

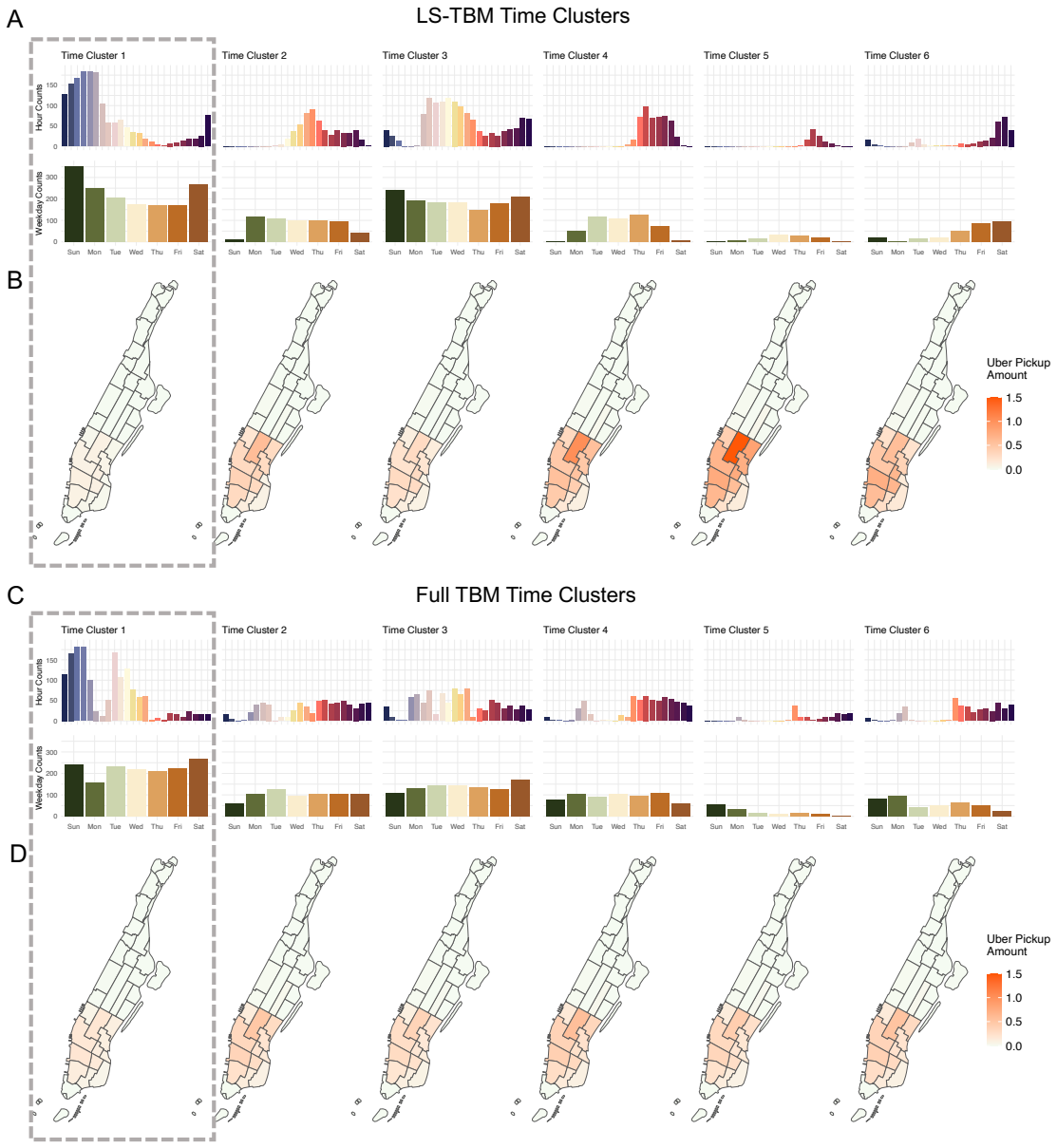

Figure C.3: Time cluster results of LS-TBM with $m_1 = 0.1n_1, (m_2, m_3) = 0.5(n_2, n_3)$ and TBM for Uber Pickup data. TBM Time Clusters are relabeled for better comparison. Subplots **A** and **C** report the counts of 24 hours and 7 weekdays for the timestamps in each time cluster. In the hour barplots, the x-axis ranges from 0 to 24. Subplots **B** and **D** visualize the estimated hourly Uber pickup amount in Manhattan neighborhoods associated with 6 time clusters. For both LS-TBM and TBM results, Time Cluster 1 (gray dashed rectangles) has the largest group size and the smallest averaged pickup amount.

C.3 Proofs

In proofs, for convenience, we use a mixture of notations $\boldsymbol{A}(j,:)$ and $\boldsymbol{A}_{j:}$ to denote the j-th row of matrix $\boldsymbol{A}$. Similar notations are used for the rows of a matrix.

Proof of Theorem 4.4

Proof of Theorem 4.4. Without loss of generality, we assume that $\sigma = 1$, the permutations $\tilde{\pi}_k = \arg\min_{\pi \in \Pi_K} \sum_{i \in S_k} \mathbb{1}\{\tilde{z}_k(i) \neq \pi \circ z_k(i)\}$ are identity mapping on $[r_k]$. We focus on the derivation for the misclassification error on the first mode $k = 1$. We drop the subscript 1 for some matricizations, e.g., $\boldsymbol{Y}, \boldsymbol{C}$, without misunderstanding. The derivations for other modes can be easily extended.

The key proof idea is to decompose the misclassification loss for seeded clustering, $L(\hat{z}_1^c)$, by two parts: the unavoidable statistical loss from the noise, and the inherited loss from the imperfect seeds. Under the model (4.1), we are able to upper bound the statistical error by the concentration properties of sub-Gaussian variables. For the inherited loss, intuitively, more accurate seeds will lead to less mistakes in the classifications for the rest of nodes. Our main goal is to quantify both statistical and inherited losses to upper bound $L(\hat{z}_1^c)$, and thereof finally obtain the error bound for $\ell(z_1^c)$ with Lemma C.1.

We first introduce extra notations for the proof.

- Complement set for seeds:

$$S_k^c = [n_k]/S_k, \quad k \in [K].$$

- Normalized membership matrices for subsets S_k's:

$$\boldsymbol{W}_k \coloneqq \boldsymbol{M}_k(S_k,:)(\text{diag}(\mathbf{1}_{m_k}^T \boldsymbol{M}_k(S_k,:)))^{-1}, \quad \tilde{\boldsymbol{W}}_k \coloneqq \tilde{\boldsymbol{M}}_k(S_k,:)(\text{diag}(\mathbf{1}_{m_k}^T \tilde{\boldsymbol{M}}_k(S_k,:)))^{-1},$$

  where $\tilde{\boldsymbol{M}}_k(S_k,:)$ are membership matrices corresponding to the seed $\tilde{z}_k$, for $k \in [K]$.

- Dual normalized membership matrices:

$$\boldsymbol{V} \coloneqq \boldsymbol{W}_2 \otimes \cdots \otimes \boldsymbol{W}_K, \quad \tilde{\boldsymbol{V}} \coloneqq \tilde{\boldsymbol{W}}_2 \otimes \cdots \otimes \tilde{\boldsymbol{W}}_K,$$

  where $\otimes$ refers to the matrix Kronecker product.

- Sub-tensors corresponding to the subsets S_k's and their matricizations:

$$\begin{aligned} &\mathcal{Y}_S \coloneqq \mathcal{Y}(S_1, \ldots, S_K), && \mathcal{Y}_{j,S_{-1}} \coloneqq \mathcal{Y}(j, S_2, \ldots, S_K) \\ &\boldsymbol{Y}_S \coloneqq \mathrm{Mat}_1(\mathcal{Y}_S), && \boldsymbol{Y}_{j,S_{-1}} \coloneqq \mathrm{Mat}_1(\mathcal{Y}_{j,S_{-1}}). \end{aligned}$$

  Similar notations are also defined for $\mathcal{X}, \mathcal{E}$.

- Estimation of core tensor with seeds and oracle estimator given true assignments:

$$\tilde{\mathcal{C}} = \mathcal{Y}_S \times_1 \tilde{\boldsymbol{W}}_1^T \times_2 \cdots \times_K \tilde{\boldsymbol{W}}_K^T, \quad \bar{\mathcal{C}} = \mathcal{Y}_S \times_1 \boldsymbol{W}_1^T \times_2 \cdots \times_K \boldsymbol{W}_K^T.$$

  We use matricizations $\tilde{\boldsymbol{C}} = \mathrm{Mat}_1(\tilde{\mathcal{C}})$, $\bar{\boldsymbol{C}} = \mathrm{Mat}_1(\bar{\mathcal{C}})$, and $\boldsymbol{C} = \mathrm{Mat}_1(\mathcal{C})$.

Next, we **decompose** the misclassification loss $L(z_1^c)$ into two parts. Consider an arbitrary node $j \in S_1^c$ with true assignment $z_1(j)$. We consider the key event in which node j is misclassified to cluster $b \neq z_1(j)$:

$$\hat{z}_1(j) = b \quad \Leftrightarrow \quad \|\tilde{\boldsymbol{C}}_{b:} - \boldsymbol{A}_{j:}\|^2 \leq \|\tilde{\boldsymbol{C}}_{z_1(j):} - \boldsymbol{A}_{j:}\|^2, \tag{C.1}$$

where $\boldsymbol{A}_{j:} = \boldsymbol{Y}_{j,S_{-1}} \tilde{\boldsymbol{V}}$ by the Line 8 in Algorithm 4.

After re-arrangement, the event (C.1) is equivalent to

$$2 \left\langle \boldsymbol{E}_{j,S_{-1}} \boldsymbol{V}, \bar{\boldsymbol{C}}_{z_1(j):} - \bar{\boldsymbol{C}}_{b:} \right\rangle \leq -\|\boldsymbol{C}_{z_1(j):} - \boldsymbol{C}_{b:}\|^2 + F_{jb} + G_{jb} + H_{jb},$$

where

$$F_{jb} = 2\left\langle \boldsymbol{E}_{j,S_{-1}}\tilde{\boldsymbol{V}}, (\bar{\boldsymbol{C}}_{z_1(j)\cdot} - \tilde{\boldsymbol{C}}_{z_1(j)\cdot}) - (\bar{\boldsymbol{C}}_{b\cdot} - \tilde{\boldsymbol{C}}_{b\cdot})\right\rangle + 2\left\langle \boldsymbol{E}_{j,S_{-1}}(\boldsymbol{V} - \tilde{\boldsymbol{V}}), \bar{\boldsymbol{C}}_{z_1(j)\cdot} - \bar{\boldsymbol{C}}_{b\cdot}\right\rangle$$

$$\begin{aligned} G_{jb} &= \left(\|\boldsymbol{X}_{j,S_{-1}}\tilde{\boldsymbol{V}} - \tilde{\boldsymbol{C}}_{z_1(j)\cdot}\|^2 - \|\boldsymbol{X}_{j,S_{-1}}\tilde{\boldsymbol{V}} - \boldsymbol{W}_{1,:a}^T\boldsymbol{Y}_S\tilde{\boldsymbol{V}}\|^2\right) \\ &\quad - \left(\|\boldsymbol{X}_{j,S_{-1}}\tilde{\boldsymbol{V}} - \tilde{\boldsymbol{C}}_{b\cdot}\|^2 - \|\boldsymbol{X}_{j,S_{-1}}\tilde{\boldsymbol{V}} - \boldsymbol{W}_{1,:b}^T\boldsymbol{Y}_S\tilde{\boldsymbol{V}}\|^2\right) \\ H_{jb} &= \|\boldsymbol{X}_{j,S_{-1}}\tilde{\boldsymbol{V}} - \boldsymbol{W}_{1,:a}^T\boldsymbol{Y}_S\tilde{\boldsymbol{V}}\|^2 - \|\boldsymbol{X}_{j,S_{-1}}\tilde{\boldsymbol{V}} - \boldsymbol{W}_{1,:b}^T\boldsymbol{Y}_S\tilde{\boldsymbol{V}}\|^2 + \|\boldsymbol{C}_{z_1(j)\cdot} - \boldsymbol{C}_{b\cdot}\|^2. \end{aligned}$$

Then, we are able to upper bound the indicator as

$$\begin{aligned} \mathbb{1}\{\hat{z}_1(j) = b\} \le \mathbb{1}\{\hat{z}_1(j) = b, \left\langle \boldsymbol{E}_{j,S_{-1}}\boldsymbol{V}, \bar{\boldsymbol{C}}_{z_1(j)\cdot} - \bar{\boldsymbol{C}}_{b\cdot}\right\rangle \le -\frac{1}{4}\|\boldsymbol{C}_{z_1(j)\cdot} - \boldsymbol{C}_{b\cdot}\|^2\} \\ + \mathbb{1}\{\hat{z}_1(j) = b, \frac{1}{2}\|\boldsymbol{C}_{z_1(j)\cdot} - \boldsymbol{C}_{b\cdot}\|^2 \le F_{jb} + G_{jb} + H_{jb}\}. \end{aligned}$$

Further, we can upper bound the misclassification loss

$$L(\hat{z}_1^c) \le \xi_1 + \frac{1}{n_1 - m_1}\sum_{j\in S_1^c}\sum_{b\in[r_1]}\zeta_{jb},$$

where

$$\begin{aligned} \xi_1 &= \frac{1}{n_1 - m_1}\sum_{j\in S_1^c}\sum_{b\in[r_1]}\mathbb{1}\{\hat{z}_1(j) = b, \left\langle \boldsymbol{E}_{j,S_{-1}}\boldsymbol{V}, \bar{\boldsymbol{C}}_{z_1(j)\cdot} - \bar{\boldsymbol{C}}_{b\cdot}\right\rangle \\ &\le -\frac{1}{4}\|\boldsymbol{C}_{z_1(j)\cdot} - \boldsymbol{C}_{b\cdot}\|^2\}\|\boldsymbol{C}_{z_1(j)\cdot} - \boldsymbol{C}_{b\cdot}\|^2, \end{aligned}$$

and

$$\zeta_{jb} = \mathbb{1}\{\hat{z}_1(j) = b, \frac{1}{2}\|\boldsymbol{C}_{z_1(j)\cdot} - \boldsymbol{C}_{b\cdot}\|^2 \le F_{jb} + G_{jb} + H_{jb}\}\|\boldsymbol{C}_{z_1(j)\cdot} - \boldsymbol{C}_{b\cdot}\|^2.$$

Here, ξ_1 is the **statistical loss** due to the existence of noise $\mathcal{E}$. In contrast, the term ζ_{jb} is controlled by the difference between the oracle estimator $\bar{\mathcal{C}}$ and the estimator based on seeds $\tilde{\mathcal{C}}$, which relies on the accuracy of seeds. The sum of ζ_{jb} is

the **inherited loss**, and should be bounded by the losses of seeds $L(\tilde{z}_k)$'s.

Indeed, Lemmas C.6 and C.5 provide the upper bounds of ξ_1 and the sum of ζ_{jb}. We have

$$L(\hat{z}_1^c) \leq \exp\left(-\frac{c_2 m_{-1}}{r_{-1}}\Delta_{\min}^2\right) + c_1 L(\hat{z}_1^c) + cC_1\sqrt{\bar{m}}\ell(\hat{z}_1^c)\sum_{k\in[K]} L(\tilde{z}_k), \tag{C.2}$$

with high probability $1-\exp(-c_3 m_1) - \exp(-c_4 m_{-1}\Delta_{\min}^2)$ and positive constants c in condition (4.3), $c_1 \in (0,1)$, and $c_2, c_3, C_1 \in \mathbb{R}_+$.

By Lemma C.1, we have $\ell(z_1^c) \leq L(z_1^c)/\Delta_{\min}^2$. Dividing $\Delta_{\min}^2$ on both sides of the inequality (C.2) and rearranging the terms, we have

$$\begin{aligned}\ell(\hat{z}_1^c) \leq \frac{L(\hat{z}_1^c)}{\Delta_{\min}^2} &\leq C_2\Delta_{\min}^{-2}\exp\left(-\frac{c_2 m_{-1}}{r_{-1}}\Delta_{\min}^2\right) + cC_1\ell(\hat{z}_1^c)\frac{\sqrt{\bar{m}}\sum_{k\in[K]} L(\tilde{z}_k)}{\Delta_{\min}^2} \\ &\leq C_2\Delta_{\min}^{-2}\exp\left(-\frac{c_2 m_{-1}}{r_{-1}}\Delta_{\min}^2\right) + Kc^2C_1\ell(\hat{z}_1^c).\end{aligned}$$

When c is small enough such that $Kc^2C_1 < 1$, we finally have

$$\ell(\hat{z}_1^c) \leq C'\Delta_{\min}^{-2}\exp\left(-\frac{c' m_{-1}}{r_{-1}}\Delta_{\min}^2\right),$$

with probability at least $1-\exp(-c''\underline{m}) - \exp(-c'''m_{-1}\Delta_{\min}^2)$.

□

Useful lemmas for the proof of Theorem 4.4

In this section, we first list some intermediate lemmas (Lemmas C.2, C.3, C.4) for sub-Gaussian concentration properties and secondary conclusions for misclassification loss analysis. Then, we state and prove the main lemmas (Lemmas C.6 and C.5) for the upper bounds of statistical and inherited losses.

The intermediate lemmas are similar with the lemmas for Han et al. (2022a, Proof of Theorem 2) but with different dimension and subject considerations. For clearness, we only show the intermediate results directly used in Lemmas C.6 and

C.5 and highlight the necessary adjustments in the proof. Full lemma statements and proofs can be found in Han et al. (2022a, Proof of Theorem 2).

Recall that we use $C, C_0, C_1, \ldots$ and $c, c_0, c_1, \ldots$ for generic large and small positive constants, respectively.

Lemma C.1 (Misclassification error and loss)**.** *Consider the misclassification error and loss for $\hat{z}_k$ under the TBM* (4.1)*. We have*

$$\ell(\hat{z}_k) \leq L(\hat{z}_k)/\Delta_{\min}^2, \quad k \in [K].$$

Lemma C.2 (Sub-Gaussian concentration)**.** *Under the set up of Theorem 4.4, as $m_k \to \infty, k \in [K]$, for any $j \in S_1^c$, we have*

$$\|\boldsymbol{E}_{j,S_{-1}}\boldsymbol{V}\| \lesssim \sqrt{\frac{r_{-1}}{m_{-1}}}(1+\sqrt{r_{-1}}). \tag{C.3}$$

Proof of Lemma C.2. Lemma C.2 is the Corollary of Han et al. (2022a, Lemma 9). □

Lemma C.3 (Upper bound for membership matrix difference)**.** *Under the set up of Theorem 4.4 and Lemma C.2, as $m_k \to \infty, k \in [K]$, we have*

$$\|\boldsymbol{E}_{j,S_{-1}}(\boldsymbol{V}-\tilde{\boldsymbol{V}})\| \lesssim \sqrt{\frac{r_{-1}(r_{-1}+\bar{m}\bar{r})}{m_{-1}}} \sum_{k=2}^{K} \frac{r_k L(\tilde{z}_k)}{\Delta_k^2}.$$

Proof of Lemma C.3. See Han et al. (2022a, Lemma 5). □

Lemma C.4 (Upper bounds for F_{jb}, G_{jb}, H_{jb})**.** *Under the set up of Theorem 4.4, as $m_k \to \infty, k \in [K]$, we have*

$$\max_{j\in S_1^c}\max_{b\neq z_1(j)}\frac{F_{jb}^2}{\|\boldsymbol{C}_{z_1(j):}-\boldsymbol{C}_{b:}\|^2}\lesssim\frac{\sum_{k=1}^K r_kL(\tilde{z}_k)}{\Delta_1^2}\|\boldsymbol{E}_{j,S_{-1}}\boldsymbol{V}\|^2 + \left(1+\frac{\sum_{k=1}^K r_kL(\tilde{z}_k)}{\Delta_1^2}\right)\|\boldsymbol{E}_{j,S_{-1}}(\boldsymbol{V}-\tilde{\boldsymbol{V}})\|^2, \tag{C.4}$$

$$\max_{j\in S_1^c}\max_{b\neq z_1(j)}\frac{G_{jb}^2}{\|\boldsymbol{C}_{z_1(j):}-\boldsymbol{C}_{b:}\|^2}\leq c_0\left(\Delta_1^2+\sum_{k=1}^K L(\tilde{z}_k)\right), \tag{C.5}$$

$$\max_{j\in S_1^c}\max_{b\neq z_1(j)}\frac{|H_{jb}|}{\|\boldsymbol{C}_{z_1(j):}-\boldsymbol{C}_{b:}\|^2}\leq\frac{1}{4}, \tag{C.6}$$

where $c_0<\frac{1}{128K}$ is some small constant.

Proof of Lemma C.4. Notice that by Assumption 4.3, there exists some $i\in S_1$ such that $z_1(i)=z_1(j)$ and thus $\boldsymbol{X}_{j,S_{-1}}=\boldsymbol{X}_{i,S_{-1}}$. Therefore, we are able to upper bound F_{jb},G_{jb},H_{jb} following the proofs of inequalities (75), (76), and (77) in Han et al. (2022a). □

Lemma C.5 (Upper bound for the sum of ζ_{jb})**.** *Under the set up of Theorem 4.4, we have*

$$\frac{1}{n_1-m_1}\sum_{j\in S_1^c}\sum_{b\in[r_1]/z_1(j)}\zeta_{jb}\leq c_1L(\hat{z}_1^c)+cC_1\sqrt{m}\ell(\hat{z}_1^c)\sum_{k\in[K]}L(\tilde{z}_k)$$

for some constants $c_1\in(0,1)$ and $C_1>0$, and c is the small constant in condition (4.3)*.*

Proof of Lemma C.5. Note that

$$\begin{aligned}\zeta_{jb}&=\mathbb{1}\{\hat{z}_1(j)=b,\frac{1}{2}\|\boldsymbol{C}_{z_1(j):}-\boldsymbol{C}_{b:}\|^2\leq F_{jb}+G_{jb}+H_{jb}\}\|\boldsymbol{C}_{z_1(j):}-\boldsymbol{C}_{b:}\|^2\\&\leq\mathbb{1}\{\hat{z}_1(j)=b,\frac{1}{4}\|\boldsymbol{C}_{z_1(j):}-\boldsymbol{C}_{b:}\|^2\leq F_{jb}+G_{jb}\}\|\boldsymbol{C}_{z_1(j):}-\boldsymbol{C}_{b:}\|^2\\&\leq 64\mathbb{1}\{\hat{z}_1(j)=b\}\left(\frac{F_{jb}^2}{\|\boldsymbol{C}_{z_1(j):}-\boldsymbol{C}_{b:}\|^2}+\frac{G_{jb}^2}{\|\boldsymbol{C}_{z_1(j):}-\boldsymbol{C}_{b:}\|^2}\right),\end{aligned}$$

where the first inequality follows from inequality (C.6) in Lemma C.4, and the last inequality follows by the fact that $\mathbb{1}\{1\leq x\}\leq x^2$ for $x\in\mathbb{R}$.

We first consider the summation over F_{jb}^2. With inequality (C.4) in Lemma C.4, we have

$$\begin{aligned}
&\frac{1}{n_1 - m_1} \sum_{j \in S_1^c} \sum_{b \in [r_1]/z_1(j)} \mathbb{1}\{\hat{z}_1(j) = b\} \frac{F_{jb}^2}{\|\boldsymbol{C}_{z_1(j):} - \boldsymbol{C}_{b:}\|^2} \\
\leq &\frac{1}{n_1 - m_1} \sum_{j \in S_1^c} \mathbb{1}\{\hat{z}_1(j) \neq z_1(j)\} \max_{b \neq z_1(j)} \frac{F_{jb}^2}{\|\boldsymbol{C}_{z_1(j):} - \boldsymbol{C}_{b:}\|^2} \\
\leq &F_1 + F_2, \qquad \text{(C.7)}
\end{aligned}$$

where

$$\begin{aligned}
F_1 &= \frac{1}{n_1 - m_1} \sum_{j \in S_1^c} \mathbb{1}\{\hat{z}_1(j) \neq z_1(j)\} \frac{\sum_{k=1}^K r_k L(\tilde{z}_k)}{\Delta_1^2} \|\boldsymbol{E}_{j,S_{-1}} \boldsymbol{V}\|^2, \\
F_2 &= \frac{1}{n_1 - m_1} \sum_{j \in S_1^c} \mathbb{1}\{\hat{z}_1(j) \neq z_1(j)\} \left(1 + \frac{\sum_{k=1}^K r_k L(\tilde{z}_k)}{\Delta_1^2}\right) \|\boldsymbol{E}_{j,S_{-1}} (\boldsymbol{V} - \tilde{\boldsymbol{V}})\|^2.
\end{aligned}$$

By inequality (C.3) in Lemma C.2, we have

$$F_1 \lesssim \frac{1}{n_1 - m_1} \sum_{j \in S_1^c} \mathbb{1}\{\hat{z}_1(j) \neq z_1(j)\} \frac{\sum_{k=1}^K r_k L(\tilde{z}_k)}{\Delta_{\min}^2} \frac{r_{-1}^2}{m_{-1}} \lesssim \ell(\hat{z}_1^c) \sum_{k \in [K]} L(\tilde{z}_k), \qquad \text{(C.8)}$$

given condition (4.3) such that $\Delta_{\min}^2 \gtrsim m_{-1}^{-1}$.

By Lemma C.3, we have

$$\begin{aligned}
F_2 &\lesssim \frac{1}{n_1 - m_1} \sum_{j \in S_1^c} \mathbb{1}\{\hat{z}_1(j) \neq z_1(j)\} \left(1 + \frac{\sum_{k=1}^K r_k L(\tilde{z}_k)}{\Delta_1^2}\right) \frac{r_{-1}^2 + \bar{m}\bar{r}}{m_{-1}} \sum_{k=2}^K \frac{r_k^2 L^2(\tilde{z}_k)}{\Delta_k^4} \\
&\lesssim \frac{1}{n_1 - m_1} \sum_{j \in S_1^c} \mathbb{1}\{\hat{z}_1(j) \neq z_1(j)\} \frac{\sqrt{\bar{m}}}{m_{-1}} \frac{\sum_{k=1}^K L(\tilde{z}_k)}{\Delta_{\min}^2} \\
&\lesssim \sqrt{\bar{m}} \ell(\hat{z}_1^c) \frac{\sum_{k=1}^K L(\tilde{z}_k)}{m_{-1} \Delta_{\min}^2} \\
&\leq c C_1 \sqrt{\bar{m}} \ell(\hat{z}_1^c) \sum_{k=1}^K L(\tilde{z}_k), \qquad \text{(C.9)}
\end{aligned}$$

where the second inequality follows by condition (4.3) such that $L(\tilde{z}_k)/\Delta_{\min}^2 \leq c/\sqrt{\bar{m}}$ for all $k \in [K]$ and the last inequality follows by $\Delta_{\min}^2 \gtrsim m_{-1}^{-1}$.

Next, we consider the summation over G_{jb}^2. With inequality (C.5) in Lemma C.4, we have

$$\begin{aligned}
&\frac{1}{n_1 - m_1} \sum_{j \in S_1^c} \sum_{b \in [r_1]/z_1(j)} \mathbb{1}\{\hat{z}_1(j) = b\} \frac{G_{jb}^2}{\|\boldsymbol{C}_{z_1(j):} - \boldsymbol{C}_{b:}\|^2} \\
\leq &\frac{1}{n_1 - m_1} \sum_{j \in S_1^c} \mathbb{1}\{\hat{z}_1(j) \neq z_1(j)\} \max_{b \neq z_1(j)} \frac{G_{jb}^2}{\|\boldsymbol{C}_{z_1(j):} - \boldsymbol{C}_{b:}\|^2} \\
\leq &\frac{c_0}{n_1 - m_1} \sum_{j \in S_1^c} \mathbb{1}\{\hat{z}_1(j) \neq z_1(j)\} \left(\Delta_1^2 + \sum_{k=1}^K L(\tilde{z}_k) \right) \\
\leq &c_1 L(\hat{z}_1^c) + c_0 \ell(\hat{z}_1^c) \sum_{k=1}^K L(\tilde{z}_k), \qquad \text{(C.10)}
\end{aligned}$$

for some $c_1 \in (0, 1)$.

Taking C_1 large enough, we have term F_2 dominants F_1 and the second term in inequality (C.10). Then, combining the upper bounds (C.7), (C.8), (C.9), and (C.10), we have shown the upper bound for the sum of ζ_{jb} in Lemma C.5. □

Lemma C.6 (Upper bound for ξ_1)**.** *Under the set up of Theorem 4.4, we have*

$$\xi_1 \leq \exp\left(-\frac{c_2 m_{-1}}{r_{-1}} \Delta_{\min}^2 \right)$$

with probability at least $1 - \exp(-c_3 m_1) - \exp(-c_4 m_{-1} \Delta_{\min}^2)$ *for some small positive constants* c_2, c_3, c_4.

Proof of Lemma C.6. Recall that

$$\mathbb{E}[\xi_1] = \frac{1}{n_1 - m_1} \sum_{j \in S_1^c} \sum_{b \in [r_1]/z_1(j)} \|\boldsymbol{C}_{z_1(j):} - \boldsymbol{C}_{b:}\|^2 \mathbb{P}(\left\langle e_j, \bar{\boldsymbol{C}}_{z_1(j):} - \bar{\boldsymbol{C}}_{b:} \right\rangle \leq -\frac{1}{4} \|\boldsymbol{C}_{z_1(j):} - \boldsymbol{C}_{b:}\|^2)$$

where $e_j := \boldsymbol{E}_{j,S_{-1}\boldsymbol{V}}$. Note that e_j's are independent random vectors in $\mathbb{R}^{r_{-1}}$, whose entries are independently sub-Gaussian distributed with norm bounded by $\mathcal{O}(\sqrt{r_{-1}/m_{-1}})$ based on Lemma C.2.

We have the upper bound for the probability

$$\begin{aligned}
&\mathbb{P}(\Big\langle e_j, \bar{\boldsymbol{C}}_{z_1(j)\colon} - \bar{\boldsymbol{C}}_{b\colon}\Big\rangle \le -\frac{1}{4}\|\boldsymbol{C}_{z_1(j)\colon} - \boldsymbol{C}_{b\colon}\|^2) \\
&\le \mathbb{P}(\Big\langle e_j, \boldsymbol{C}_{z_1(j)\colon} - \boldsymbol{C}_{b\colon}\Big\rangle \le -\frac{1}{8}\|\boldsymbol{C}_{z_1(j)\colon} - \boldsymbol{C}_{b\colon}\|^2) \\
&\qquad + \mathbb{P}(\Big\langle e_j, \bar{\boldsymbol{C}}_{z_1(j)\colon} - \boldsymbol{C}_{z_1(j)\colon}\Big\rangle \le -\frac{1}{16}\|\boldsymbol{C}_{z_1(j)\colon} - \boldsymbol{C}_{b\colon}\|^2) \\
&\qquad + \mathbb{P}(\Big\langle e_j, \boldsymbol{C}_{b\colon} - \bar{\boldsymbol{C}}_{b\colon}\Big\rangle \le -\frac{1}{16}\|\boldsymbol{C}_{z_1(j)\colon} - \boldsymbol{C}_{b\colon}\|^2) \\
&\le 5\exp\left(-\frac{cm_{-1}}{r_{-1}}\|\boldsymbol{C}_{z_1(j)\colon} - \boldsymbol{C}_{b\colon}\|^2\right), \qquad \text{(C.11)}
\end{aligned}$$

with probability $1-\exp(-c_3 m_1)$, where the last inequality follows from Han et al. (2022a, Lemma 6), replacing $\boldsymbol{S}$ terms by $\boldsymbol{C}$ terms in our context. Note that unlike Han et al. (2022a, Lemma 6), e_j is independent with $\boldsymbol{C}_{b\colon} - \bar{\boldsymbol{C}}_{b\colon}$ for all $j \in S_1^c, b \in [r_1]$ since the randomness are from two non-overlapped parts of the noise tensor. The inner produce $\Big\langle e_j, \bar{\boldsymbol{C}}_{z_1(j)\colon} - \boldsymbol{C}_{z_1(j)\colon}\Big\rangle$ turns out to be the weighted sum of products of two independent random vectors $e_j^T e_l$ for $j \in S_1^c, l \in S_1$. Since Han et al. (2022a, Lemma 6) merely relies on the Bernstein inequality for $e_j^T e_l$ with $j \neq l$, Lemma 6 also works for our case and inequality (C.11) holds.

Hence, we have upper bound of the expectation

$$\begin{aligned}
\mathbb{E}[\xi_1] &\le \frac{5}{n_1 - m_1}\sum_{j\in S_1^c}\sum_{b\in[r_1]/z_1(j)}\|\boldsymbol{C}_{z_1(j)\colon} - \boldsymbol{C}_{b\colon}\|^2\exp\left(-\frac{cm_{-1}}{r_{-1}}\|\boldsymbol{C}_{z_1(j)\colon} - \boldsymbol{C}_{b\colon}\|^2\right) \\
&\le \exp\left(-\frac{cm_{-1}}{2r_{-1}}\|\boldsymbol{C}_{z_1(j)\colon} - \boldsymbol{C}_{b\colon}\|^2\right) \\
&\le \exp\left(-\frac{cm_{-1}}{2r_{-1}}\Delta_{\min}^2\right).
\end{aligned}$$

With Markov inequality, we have

$$\mathbb{P}\left(\xi_1 \leq \mathbb{E}[\xi_1] + \exp\left(-\frac{cm_{-1}}{4r_{-1}}\Delta^2_{\min}\right)\right) \geq 1 - \exp\left(-\frac{c_3 m_{-1}}{4r_{-1}}\Delta^2_{\min}\right).$$

Then, we have finished the proof of Lemma C.6.

□

Proof of Theorem 4.5

Proof of Theorem 4.5. Without loss of generality, we only prove the sub-tensor balance property on the first mode $k = 1$. Then, we drop the subscript 1 from n_1, m_1, r_1, S_1, z_1. The balance property on other modes can be proved by modifying the subscripts.

We first define the number of nodes in a-th community and the ratio of nodes in a-th community:

$$n_a = \sum_{i\in[n]} \mathbb{1}\{z(i) = a\}, \quad m_a = \sum_{i\in S} \mathbb{1}\{z(i) = a\}, \quad \mu_a = n_a/n.$$

By random sampling, m_a follows Hypergeometric distribution with parameters (m, n_a, n) for all $a \in [r]$. Notice that m_a's are not independent but the conclusion for marginal distributions holds.

By Corollary 1 in Greene and Wellner (2017), for arbitrary $a \in [r]$, we have

$$\mathbb{P}(m_a \geq m\mu_a + mt) \leq \exp\left(-\frac{mt^2}{\sigma_a^2(1 - \frac{m-1}{m-1}) + \frac{t}{3}}\right) \leq \exp\left(-\frac{mt^2/2}{\mu_a + \frac{t}{3}}\right), \quad \text{for } t > 0, \tag{C.12}$$

where $\sigma_a^2 = \mu_a(1 - \mu_a)$ and the second inequality holds by $\mu_a < 1$.

Similarly, by symmetry, we have

$$\mathbb{P}(m_a \leq m\mu_a - mt) \leq \exp\left(-\frac{mt^2}{\sigma_a^2(1 - \frac{m-1}{m-1}) + \frac{t}{3}}\right) \leq \exp\left(-\frac{mt^2/2}{\mu_a + \frac{t}{3}}\right), \quad \text{for } t > 0.$$

Take $t = \frac{1}{4}\mu_a$. We have upper bound for the maxima of m_a

$$\begin{aligned}
\mathbb{P}\left(\max_{a\in[r]} m_a \le \frac{5}{4}m\max_{a\in[r]}\mu_a\right) &= \mathbb{P}\left(\bigcap_{a\in[r]}\{m_a \le \frac{5}{4}m\max_{a\in[r]}\mu_a\}\right) \\
&\ge \mathbb{P}\left(\bigcap_{a\in[r]}\{m_a \le \frac{5}{4}m\mu_a\}\right) \\
&\ge 1 - \sum_{a\in[r]}\exp\left(-\frac{3m\mu_a}{128}\right) \\
&\ge 1 - r\exp\left(-\frac{3m\alpha_1}{128r}\right), \qquad \text{(C.13)}
\end{aligned}$$

where the second inequality follows by the union bound and tail bound (C.12), and the last inequality holds with Assumption 4.2 indicating $\frac{\alpha_1}{r} \le \mu_a \le \frac{\alpha_2}{r}$.

Similarly, we have lower bound for the minima of m_a

$$\mathbb{P}\left(\min_{a\in[r]} m_a \ge \frac{3}{4}m\min_{a\in[r]}\mu_a\right) \ge 1 - r\exp\left(-\frac{3m\alpha_1}{128r}\right). \qquad \text{(C.14)}$$

Finally, take $\alpha_3 < \frac{3\alpha_1}{4}$ and $\alpha_4 > \frac{5\alpha_2}{4}$. We finish the proof of Theorem 4.5 by combining inequalities (C.13) and (C.14):

$$\begin{aligned}
&\mathbb{P}\left(\frac{\alpha_3 m}{r} \le m_a \le \frac{\alpha_4 m}{r},\ \text{for all } a\in[r]\right) \\
&= \mathbb{P}\left(\frac{\alpha_3 m}{r} \le \min_{a\in[r]} m_a \text{ and } \max_{a\in[r]} m_a \le \frac{\alpha_4 m}{r}\right) \\
&\ge \mathbb{P}\left(\min_{a\in[r]} m_a \ge \frac{3}{4}m\min_{a\in[r]}\mu_a \text{ and } \max_{a\in[r]} m_a \le \frac{5}{4}m\max_{a\in[r]}\mu_a\right) \\
&\ge 1 - 2r\exp\left(-\frac{3m\alpha_1}{128r}\right).
\end{aligned}$$

□

D APPENDIX OF CHAPTER 5

D.1 Extra analyses of simulated data

Simulated data generation

Our simulated dataset consists of both genotypes and expressions of p genes for n samples, represented as two n-by-p matrices G and E, respectively. To simulate genotypes, we started with homozygous parents (one with genotype AA and the other one with BB) and simulated the genotypes for an F1 cross with n samples followed by an F2 intercross generation with random chromosomal crossing overs. All F1 samples have heterozygous AB genotypes for all genes. We randomly select two individuals from the n F1 samples as parents, generate two gametes from these parents, and combine the gametes to get the genotype for one F2 individual. We repeat this step n times to get n F2 samples. For each F1 gamete, we simulate one recombination event per chromosome with a probability 0.7, randomly placed along the chromosome with a uniform distribution. Due to the homozygous parents, the minor allele frequency in this study is 0.5. For extra analysis in Section D.1, we repeated the breeding procedures to simulate genotypes for F3, F4, and F5 intercross generations, to mimic the process of generating fine-mapping populations. The simulated genotype matrices have entries with values 0, 1, or 2, where 0 and 2 refer to homozygous genotypes, and 1 refers to the heterozygous genotype. We use i for the index of gene and use j for the index of sample.

Given the simulated genotype matrix G, we simulate baseline expression level (null data) for each gene in each sample, assuming no network effects. Specifically, we simulated the null expression value of the i-th gene for the j-th sample independently based on Poisson distribution with scalar parameters $(\mu, \sigma_\alpha, \sigma_\beta)$:

$$E^0_{ji} \sim_{\text{ind}} \text{Poi}(\exp(\alpha_i + \beta_j \times G_{ji})), \quad \text{where } \alpha_i \sim_{i.i.d.} N(\mu, \sigma^2_\alpha),\ \beta_i \sim_{i.i.d.} N(0, \sigma^2_\beta).$$

The parameters μ and σ^2_α control the overall mean and deviation of the expression

data, respectively, and σ_β^2 controls the degree of genetic effects on expression. Next, we constructed the gene co-expression networks, represented as symmetric matrices $M_k \in \mathbb{R}^{p\times p}$, $k = 0, 1, 2$, for three genotypes. For genotype k, the entries in M_k follow the following distribution independently with scalar parameters (d_k, δ_k):

$$M_{k,i_1,i_2} = \text{Bernoulli}(d_k) \times N(0, \delta_k^2), \quad \text{where } 1 \le i_1 \le i_2 \le p, \text{ and } M_{k,i_1,i_2} = M_{k,i_2,i_1}.$$

The parameter d_k controls the sparsity of gene-gene correlation (i.e. number of nonzero entries in M_k), and δ_k^2 controls the magnitude of nonzero entries in M_k in genotype k. The genetic-related network effects are specified by setting various values of (d_k, δ_k^2). Furthermore, to mimic the additive genetic effects, we may set the heterozygous network as the average of homozygous networks, i.e., $M_1 = (M_0+M_2)/2$. For simplicity we do not consider more complex patterns of dominance or over-/under-dominance, which may exsit in empirical co-expression networks in hybrid populations.

Last, we imposed the co-expression network effect on top of the the null expression. We randomly selected one gene, say i^*, as the snQTL. The genotype of snQTL i^* determined which matrix M_k is for the co-expression network effect. We alter the the expressions for genes involved in co-expressed pairs, using the following calculations:

$$E_{ji} = E_{ji}^0 + \sum_{l\le i} M_{G_{ji^*},li} \times (E_{ji}^0 - \exp(\alpha_l + \beta_l \times G_{jl})),$$

where the second term on the right-hand side is the network effect of snQTL i^* to the expression of gene i in sample j. It is possible for above equation to generate negative expressions, denoted as E_{ji}^N, which is impossible in real life. The simplest way is to set negative expressions as 0; however this simple way leads the undesired concentration around 0 compared with real data. Therefore, we consider the regeneration for negative expression:

$$E_{ji} \sim_{\text{ind}} \text{Poi}(\exp(\mu + \eta * E_{ji}^N)),$$

where η is a positive constant. Intuitively, if we have a negative expression with larger absolute value, we will have a low expectation in above Poisson distribution.

In our simulation, we generated data of $p = 200$ genes located on 20 chromosomes with varying population size n from 50 to 500. We considered additive network effect and tuned the parameters to mimic the distribution of expression levels from the empirical stickleback dataset, by choosing:

$$\mu = 2.6,\ \sigma_\alpha = 1,\ \sigma_\beta = 0.75,\ d_0 = 0.01,\ \delta_0 = 0.5,\ d_2 = 0.03, \delta_2 = 5,\ \eta = 0.05.$$

Raw correlation map for simulated data

Our analysis compared the absolute genetic correlation heatmaps between real stickleback data Weber et al. (2022) and synthetic data (Figure D.1). Unlike Figure 2A (real data), the synthetic heatmap shows correlations ranging from -1 to 1, while real marker pairs mostly have positive correlations. This difference arises from three aspects of our simulation design:

- Parent Genotype Initialization: In the simulation, half of first-generation diploid parents are given genotype AA while the other half are given genotype BB, for all markers. This leads to all offspring in the first generation having the same genotype AB for all markers.
- Enforced Crossover: During gamete formation, we simulate chromosomal crossover events. While these crossovers are important for real breeding, they can create negative correlations in our simulated data. This is because markers near the head and tail of the chromosome are likely to have different genotypes in the resulting offspring, even though they came from the same parent.
- Limited Breeding Generations: We only simulated breeding for two generations (F1 and F2). With more generations, these negative correlations due to initial crossovers would eventually be diluted through further recombination.

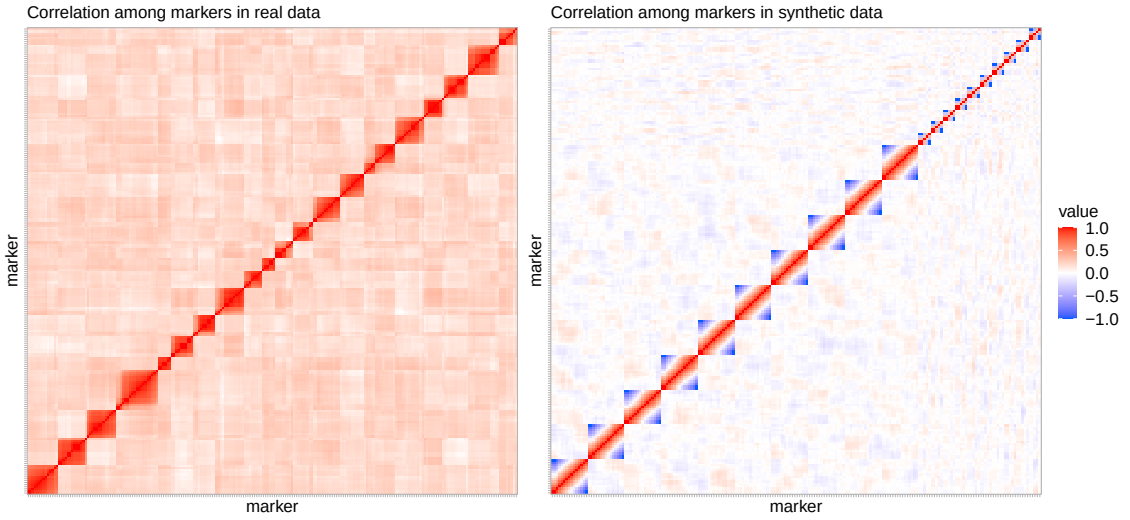

Figure D.1: Genetic correlation heatmap among markers in real F2 hybrid three-spined stickleback data and synthetic data. Blue color indicates negative correlation; red color indicates positive correlation.

For example, consider a chromosome with 5 markers, with 0 representing genotype A and 1 representing genotype B. The second-generation diploid parents have two chromatids for chromosome 1, with genotype (00000) and (11111). With a crossover at the fourth marker, possible gametes could be (00011) and (11100). Combine two random gametes to obtain the genotype of diploid offspring. If marker 1 has genotype $(0,0)$, then marker 5 must have $(1,1)$; if marker 1 has genotype $(0,1)$, then marker 5 must have $(1,0)$. Markers at opposite ends of the chromosome have opposite genotypes, leading to a negative correlation.

It is important to note that our snQTL testing is invariant to the label switch between A and B. Therefore, the presence of negative correlations in the simulated data does not impact the evaluation of our method's effectiveness.

Detailed simulation results for Fig 2C

Table D.1 shows the average p-values for all methods in simulation Figure 2C. As the population size increases, p-values for all methods decrease. However, the improvement in accuracy for the local approach is minimal compared to the rapid increase in accuracy observed with snQTL.

Method / Population Size	50	80	100	150	200	250	300	500
Max at snQTL	0.397	0.272	0.221	0.154	0.131	0.113	0.080	0.008
Tensor at snQTL	0.399	0.291	0.230	0.154	0.133	0.112	0.080	0.009
Local at snQTL	0.746	0.723	0.700	0.675	0.621	0.683	0.641	0.597
Max at non-snQTL	0.517	0.493	0.518	0.457	0.472	0.438	0.494	0.444
Tensor at non-snQTL	0.523	0.491	0.515	0.459	0.474	0.430	0.497	0.453
Local at non-snQTL	0.822	0.842	0.862	0.869	0.860	0.913	0.909	0.923

Table D.1: Averaged p-values in simulation Figure 2C.

Simulation with varying sparsity parameters

Figure D.2 presents additional simulation results with snQTL using different sparsity parameter. Following earlier works Zhu et al. (2017); Witten et al. (2009), we set $R = c^2 p$, where $c \in [0, 1]$ is the tunning parameter. Our snQTL shows a stable performance regardless the choice of sparsity. The accuracy of snQTL keeps over 0.7 for all sparsity levels. The only exception is the extremely sparse case with $c = 0.1$ letting only 1% genes contribute to the joint different network. However, the synthetic data has the true sparsity level around $c^2 = \sqrt{d_2} = \sqrt{0.03} = 17\%$ ($c = 0.42$), where d_2 is the network density in Section D.1. Combining with the main simulation in Fig 2, we conclude that our method provides a substantial and robust increase in performance with and without sparsity constraints, compared with the local method.

To test the efficacy of the data-drive BIC approach, we further consider the simulation with synthetic differential tensors $\mathcal{D}$ of different sparsity levels. Specifically,

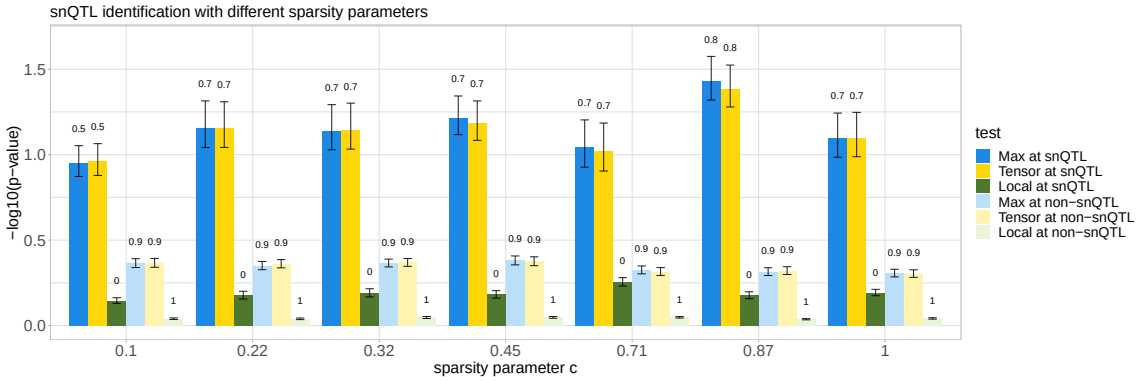

Figure D.2: Barplots comparing the snQTL identification performances for snQTL framework and local method (F-test for regression of pairwise co-expression onto genotype) on synthetic data with varying sparsity parameter $c \in [0.1, 1]$. The population size is set to 300. A larger c results in more genes contribute to the co-expression changes. True positive (or negative) rates for the tests at snQTL (or non-snQTL) are shown above the bars. All reported numbers are averaged across 50 replications for each population size.

with $p = 500$, we generate $\mathcal{D} \in \mathbb{R}^{p\times p\times 3}$ as

$$\mathcal{D} = \Lambda v^c \circ v^c \circ u + \mathcal{E},$$

where $\Lambda = 25$, v^c, u have normalized entries from Uniform$(-1, 1)$ such that $\|v^c\|_0 = c^2 p$, $\|v^c\|_2 = \|u\|_2 = 1$, and $\mathcal{E} \in \mathbb{R}^{p\times p\times 3}$ is the noise tensor with 3 symmetric p-by-p slides consisting normal distribution entries from $N(0, 0.01)$. The rough signal-to-noise ratio $\|\Lambda v^c \circ v^c \circ u\|_F / \|\mathcal{E}\|_F$ is smaller than 0.1, which mimics the small signal cases in practice. We vary the sparsity level with $c^2 = 0.05, 0.1, \ldots, 0.7$ and consider the candidates from 0.01 to 0.8. Table D.2 indicates the efficacy of our selection approach, especially for the sparse case with $c^2 \leq 0.5$. Note that in SSTD algorithm, our algorithm optimizes the objective under a relaxed convex restriction, $\|v^c\|_1 \leq c\sqrt{p}$. Thus, for SSTD outputs, c^2 is a lower bound for the proportion non-zero elements in v^c. This observation explains the BIC selection results under the relatively dense cases $c^2 > 0.5$. Though we select c's smaller than the ground

truth, the actual number of non-zero elements in v^c is larger than c^2p and well-approximate the true signal. Hence, we conclude that the proposed BIC approach is an effective for sparsity selection in principle.

True c^2	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7
Selected $\hat{c}^2$	0.05 (0)	0.1 (0)	0.2 (0)	0.3 (0)	0.4 (0)	0.5 (0)	0.51 (0.02)	0.6

Table D.2: Sparsity parameter selection with BIC approach. We vary the proportion $R = c^2p$ with $c^2 = 0.05, \ldots, 0.7$. BIC approach considers candidate c^2 from 0.01 to 0.8. For each true c^2, we repeat the simulation for 20 times and report the average and standard deviation of the selected c^2.

Simulation for synthetic data with more generations

Figure D.3 presents additional simulation results using synthetic data generated from F3, F4, and F5 hybrid crosses. These later breeding generations are often used in fine-mapping studies because they create smaller linkage windows. In essence, these later generations move closer to mimicking GWAS in genetically diverse outbred populations, where breeding is not involved. We find that our comparison conclusions remain the same to those performed with the F2 hybrid data in the main text. This consistency demonstrates that snQTL maintains its outer-performance compared to the local method, regardless of the breeding generation used in the simulation.

Simulation for GWAS-like synthetic data

We also generate GWAS-like synthetic data to show the general efficacy of snQTL. Genotype generation is the main difference between F2 and GWAS-like data. For GWAS-like synthetic data, we generate genotype with random minor allele frequency and Hardy-Weinberg distribution. Specifically, we generate the genotype for the i-th gene of the j-th sample as

$$G_{ji} \sim \text{Hardy-Weinberg}(\gamma_i), \text{ with } \gamma_i \sim \text{Uniform}[0.1, 0.9],$$

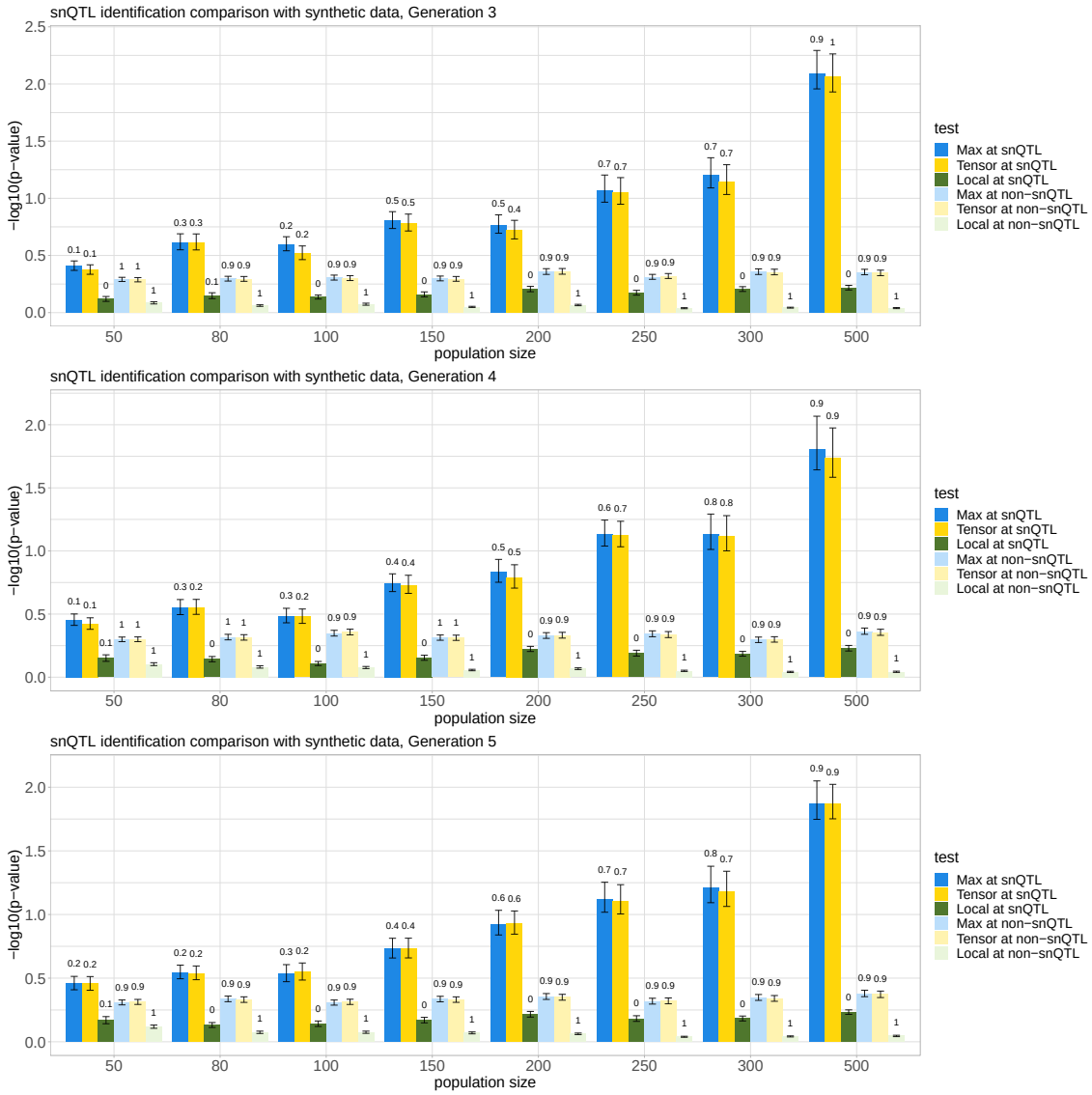

Figure D.3: Comparison between snQTL framework and local methods (F-test for regression of pairwise co-expression against genotype) on synthetic data with varying population size from 50 to 500. Synthetic datasets are generated from the F3, F4, and F5 hybrids (from top to bottom), respectively. True positive (or negative) rates for the tests at snQTL (or non-snQTL) are shown above the bars. All reported numbers are averaged across 50 replications for each population size.

independently for $i = 1, \ldots, p, j = 1, \ldots, n$. Here, for $X \sim$ Hardy-Weinberg(γ), we have

$$\mathbb{P}(X = 0) = (1 - \gamma)^2, \quad \mathbb{P}(X = 1) = \gamma(1 - \gamma), \quad \mathbb{P}(X = 2) = \gamma^2.$$

Given GWAS-like genotype, we generate the expression following the same procedure used in F2 hybrid in Section D.1.

Figure D.4 indicates that GWAS-like genotype does not show LD structure as real data, while the expression distribution of GWAS-like data in Figure D.5 is close to real data. The simulation comparsion results in Figure D.6 still supports the superior performance of our snQTL compared with the local method. Meanwhile, snQTL shows a less power than that for F2 hybrid (Figure 2). The decrease in power is expected, since GWAS genotype generation will lead more unbalanced genotype group partitions with small or large γ.

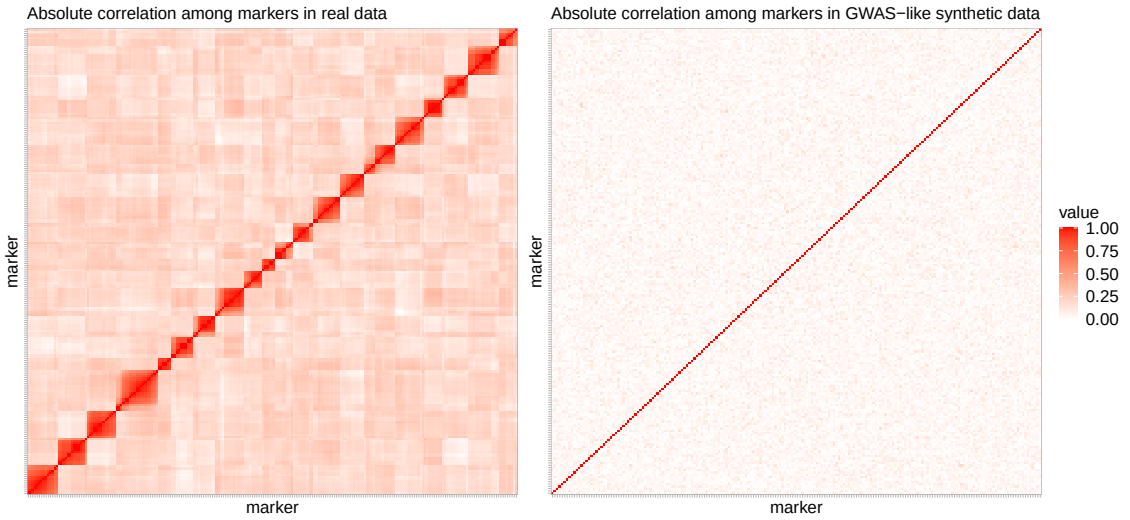

Figure D.4: Absolute genetic correlation heatmaps among the markers in real F2 hybrid three-spined stickleback data Weber et al. (2022) and GWAS-like synthetic data. Markers are ordered following their positions on the genome. Genetic correlations are measured by absolute sample Pearson correlation coefficients between the genotypes of two markers.

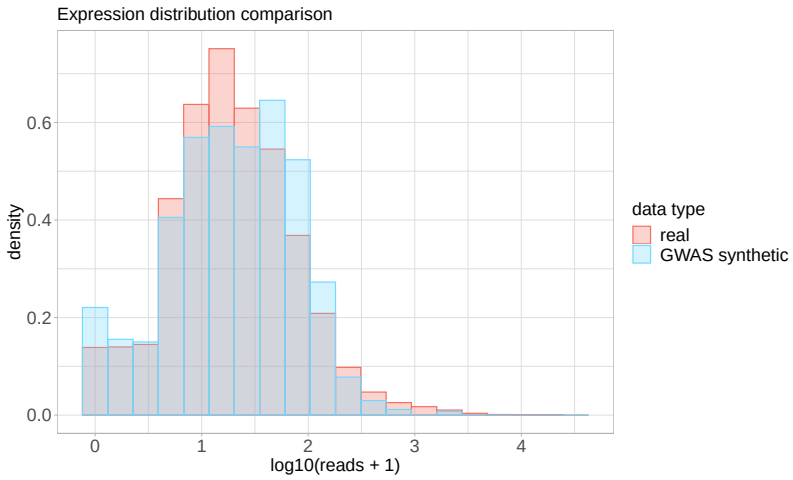

Figure D.5: Density histograms for expression counts in real stickleback and GWAS-like synthetic data. The parameters in GWAS-like synthetic data generation are the same for Fig 2 in the main text.

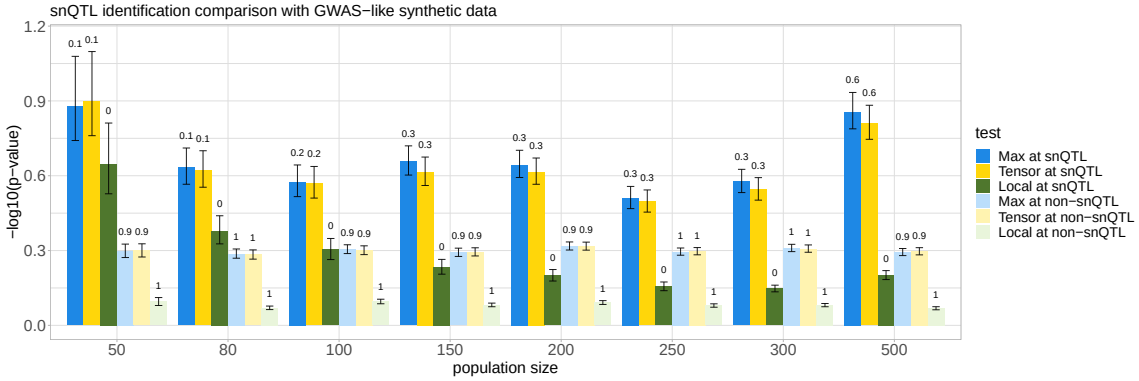

Figure D.6: Barplots comparing the snQTL identification performances for snQTL framework and local method (F-test for regression of pairwise co-expression onto genotype) on GWAS-like synthetic data with varying population size from 50 to 500. We set sparsity parameter $R = p$ in snQTL for a fair comparison with non-sparse local method. True positive (or negative) rates for the tests at snQTL (or non-snQTL) are shown above the bars. All reported numbers are averaged across 50 replications for each population size.

D.2 Pre-processing and extra analyses with stickleback data

Pre-processing of stickleback data

The original stickleback data Weber et al. (2022) includes gene expression levels (transcript counts) for 26,285 genes and genotypes for 234 genetic markers from 351 samples in the F2 and backcross generations. The data also contains covariate information such as sex, family, and infection status.

We describe our pre-processing on the gene expression matrix. Let X^0 denote the raw transcript count matrix, and $n = 351, p = 26,285, q = 234$ for the numbers of samples, genes, and markers, respectively We use i for the index of gene and use j for the index of fish.

We follow these three main steps:

(i) Normalization. We normalized the raw transcript counts to account for differences in sequencing depth across samples. This is achieved by dividing each transcript count X^0_{ji} by the total transcript count for that sample. The resulting normalized matrix is denoted by X^N with entries

$$X^N_{ji} = \frac{X^0_{ji}}{\sum_{i=1}^{p} X^0_{ji}}, \quad \text{for sample } j = 1, \ldots, n.$$

The denominator $\sum_{i=1}^{p} X^0_{ji}$ is referred to as the library size or sequencing depth of the sample j.

(ii) Covariate Removal. We removed the effects of covariates of sex and family. This was done using linear regression. The covariate sex has three levels (female, male, and non-identified), and the covariate family has 36 levels based on the parental origin. The residuals from this model represent the gene expression levels after removing the influence of these covariates.

$$X^R_{ji} = \text{residual}(X^N_{ji} \sim S_j + A_j), \quad \text{for gene } i = 1, \ldots, p.$$

Here, $y \sim x$ represents the standard repression model in which y servers as the response and x serves as predictor, and S_j denotes the sex effect and A_j denotes the family effect. We use X^R to denote the residual matrix after removing covariate effects.

As we mentioned in the main text, we also consider warm infection status as a possible covariate. We explored this further in Section D.2 using a similar regression approach with an additional binary predictor for worm presence/absence:

$$X_{ji}^R = \text{residual}(X_{ji}^N \sim S_j + A_j + W_j), \quad \text{for gene } i = 1, \ldots, p,$$

where W_j is a binary predictor encoding the worm presence/absence of the samples.

(iii) Gene Selection: Finally, we selected the top 10,000 genes with the highest adjusted mean expression m_i defined by

$$m_i = \frac{1}{n} \sum_{j=1}^{m} X_{ji}^R, \quad \text{for all genes } i = 1, \ldots, p$$

This metric represents the average expression level for each gene across all samples. We focused on highly expressed genes because they are more likely to be relevant for biological processes. The cutoff of 10,000 genes was chosen for computational efficiency.

Testing results with matrix statistics

In addition to the tensor statistics, we investigated two alternative approaches based on matrix-spectrum statistics. The first one uses the max statistic defined in the main text. The second approach utilizes a variant called the sum statistic, defined as:

$$\text{Stat}_{\text{sum}} = \lambda(D_{AB}) + \lambda(D_{AH}) + \lambda(D_{BH}),$$

where $\lambda(\cdot)$ represents the sLME, D_{AB}, D_{AH}, and D_{BH} represent pairwise differential network from the original sample correlation matrices.

As shown in Figures D.7 and D.8, we found that both the max and sum statistics identified snQTLs for sticklebacks clustered on chromosomes 3, 8, and 18. This consistency across different statistical methods strengthens the reliability of our findings.

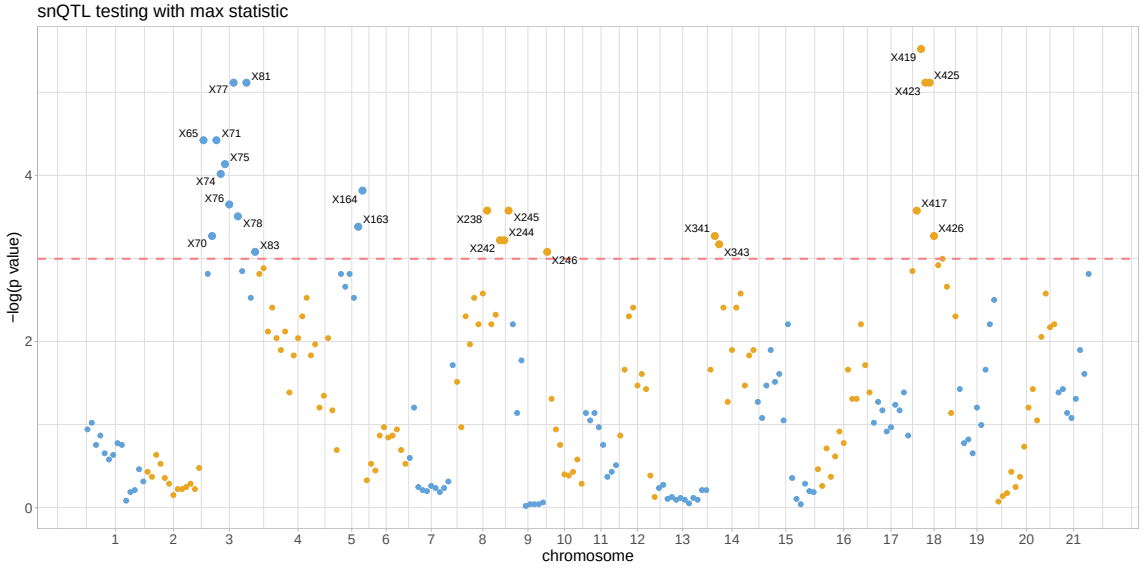

Figure D.7: Manhattan plot for snQTL testing with max statistics marks stickleback snQTLs (above pink dashed line with p-values smaller than 0.05), mainly clustered in Chr3, Chr8, and Chr18.

Testing results controlling for tape worm infection

To prioritize analyses with limited computational resources, we performed snQTL testing only on the top snQTLs identified from the non-infection-controlled expression data. We found that the top snQTLs on chromosomes 3, 8, and 18 remained the same even after controlling for infection status (Table D.3). This consistency suggests that the network effects of these snQTLs are not driven by the environ-

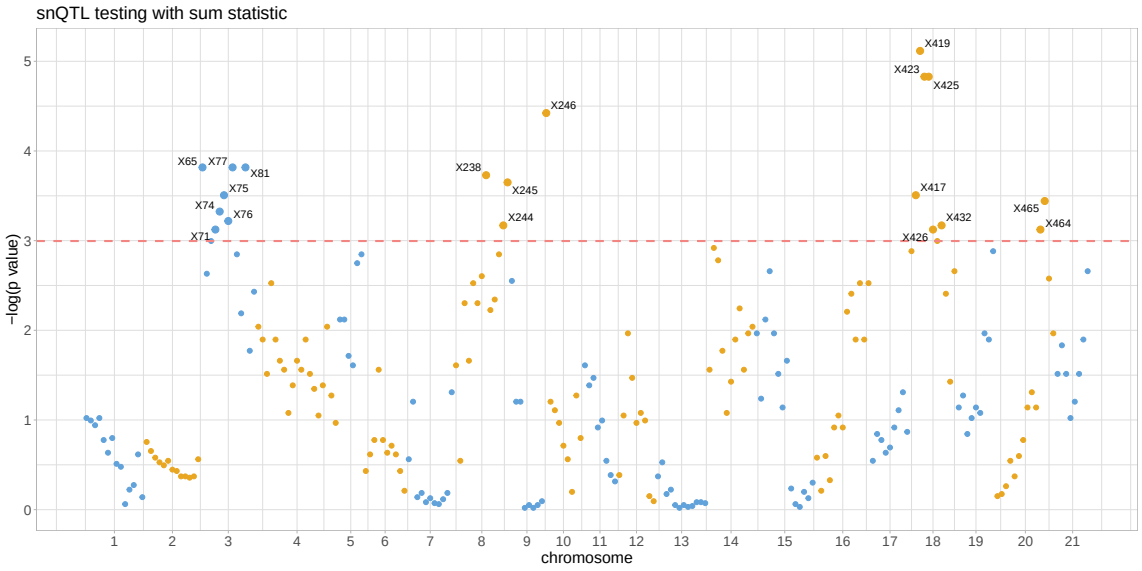

Figure D.8: Manhattan plot for snQTL testing with sum statistics marks stickleback snQTLs (above pink dashed line with p-values smaller than 0.05), mainly clustered in Chr3, Chr8, and Chr18.

mental factor of worm infection. Rerunning the analysis with infection-controlled expression data for all markers might reveal even more significant snQTLs and related discoveries. However, we will leave such additional analyses for future studies.

Markers	p-values (infection uncontrolled)	p-values (infection controlled)
X77	0.01	0.002
X81	0.012	0.004
X419	0.008	0.004
X423	0.008	0.008
X238	0.018	0.024
X245	0.02	0.016

Table D.3: Empirical p-values of top snQTLs obtained by snQTL testing with infection-uncontrolled expression (main text result) and with the infection-controlled expression (Section D.2).

Testing results excluding cross type effects

In current stickleback data pre-processing, we exclude the family effects by regressing the normalized expression onto the family covariate with 36 levels (Section D.2). In real data, the cross type information (with 3 levels, GBC, RBC, F2) for each sample is also recorded, and family covariates are nested in cross type covariates. In principle, the analysis results excluding family effects are more conservative than that excluding cross type effects. Hence, we also investigate the snQTL results only without cross type effects. Table D.2 indicates that top snQTLs on chromosomes 3, 8 remain significant while snQTLs on chromosome 18 is perturbed by different pre-processing procedures. For the most notable identified snQTL, X419, the corresponding gene leverage (Figure D.9) and joint differential network analysis without cross type effects show similar patterns as that without family effects (presented in the main text). Therefore, in this paper, we stick with the conservative family-covariate pre-processing. We leave the additional full analysis with cross-type-covariate pre-processing in the future.

Markers	p-values (without family effects)	p-values (without cross type effects)
X77	0.01	0.02
X81	0.012	0.022
X419	0.008	0.052
X423	0.008	0.044
X238	0.018	0.138
X245	0.02	0.182

Table D.4: Empirical p-values of top snQTLs obtained by snQTL testing without family effects (main text result) and without cross type effects (Section D.2).

Population branch statistic distributions on Chr 3 and Chr 8

Figures D.10 and D.11 show that several protein-coding genes lie in regions adjacent to PBS outliers near the snQTLs on Chr 3 and Chr 8.

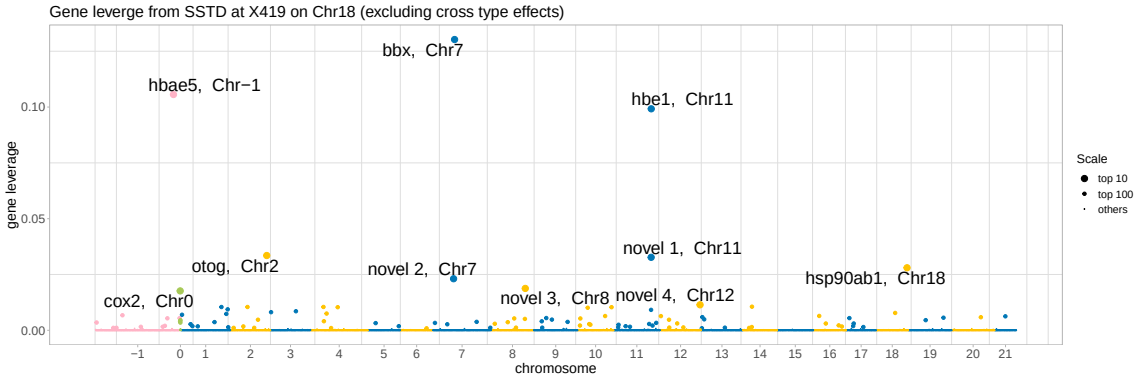

Figure D.9: Joint differential network analysis at snQTLs X419 on Chr 18 with leverage scores for 10000 genes, excluding the cross type effects. Primary genes with top 10 leverage are highlighted with transcription IDs. Mitochondrial genome (MT) and scaffold region are coded as Chr 0 and Chr -1, respectively. novel 1: ENSGACT00000018413; novel 2: ENSGACT00000026589; novel 3: ENSGACT00000017116; novel 4: ENSGACT00000017388.

Joint differential networks for snQTLs on Chr 3 and Chr 8

We further analyzed the top snQTLs identified earlier: X77 on Chr 3 and X238 on Chr 8. Interestingly, the results at these loci were highly similar to those observed at X419 on Chr 18 (see main text). Figure D.12 illustrates this similarity. The sets of top 100 genes identified by high tensor leverages at X77 and X238 show a significant overlap with the top genes for X419, especially the top 10 most important genes. However, the specific ranking of these genes might differ slightly between snQTLs. To explore this further, we constructed a joint differential network for X77 and X238, using the top genes identified at X419. Figure D.13 reveals the strong connections between primary and secondary genes. This patten is consistent across these multiple snQTLs. This consistency strengthens the evidence that our findings regarding oxygen transport pathways in the joint differential network are not simply random observations.

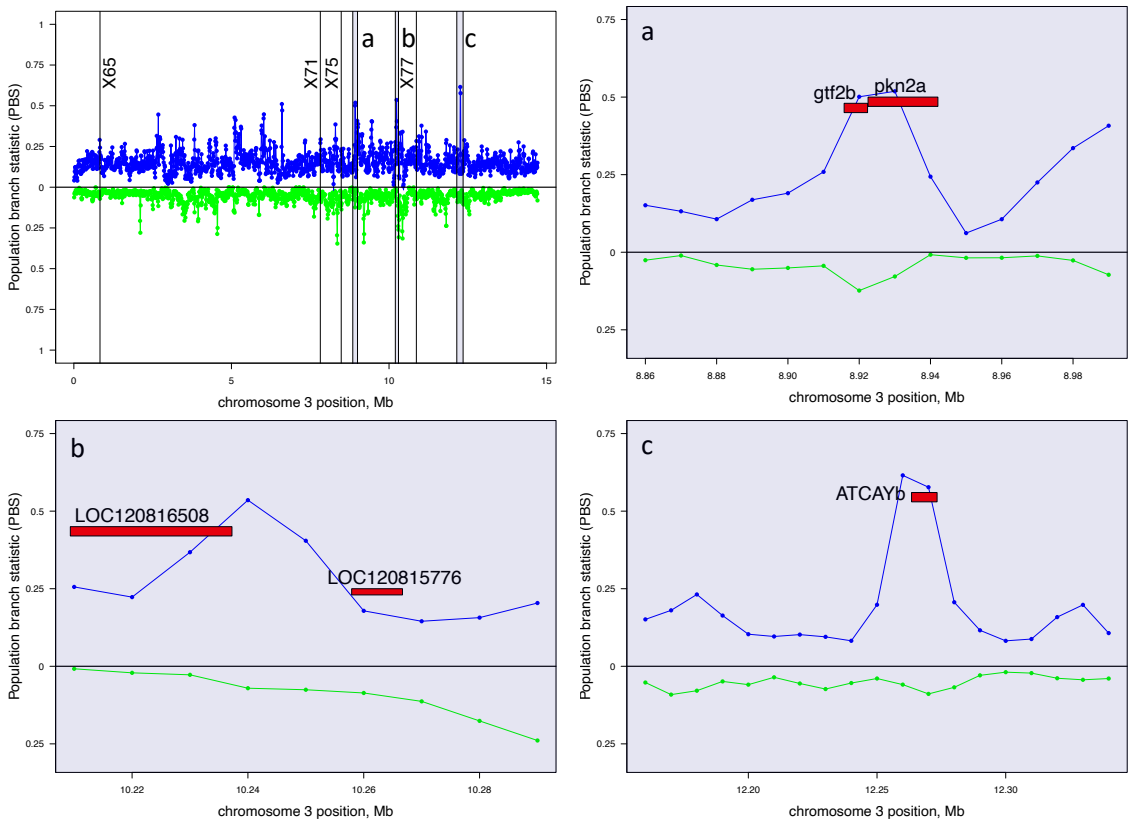

Figure D.10: Strong genomic targets of selection with high population branch statistic distribute around the outstanding snQTLs in Chr 3. Values above the medial line represent higher PBS in Gosling Lake (blue); values below the line represent higher PBS in Roberts Lake (green). Protein-coding genes lie in regions adjacent to three PBS outliers (a, b, c) around snQTLs (markers X65, X71, X75, and X77).

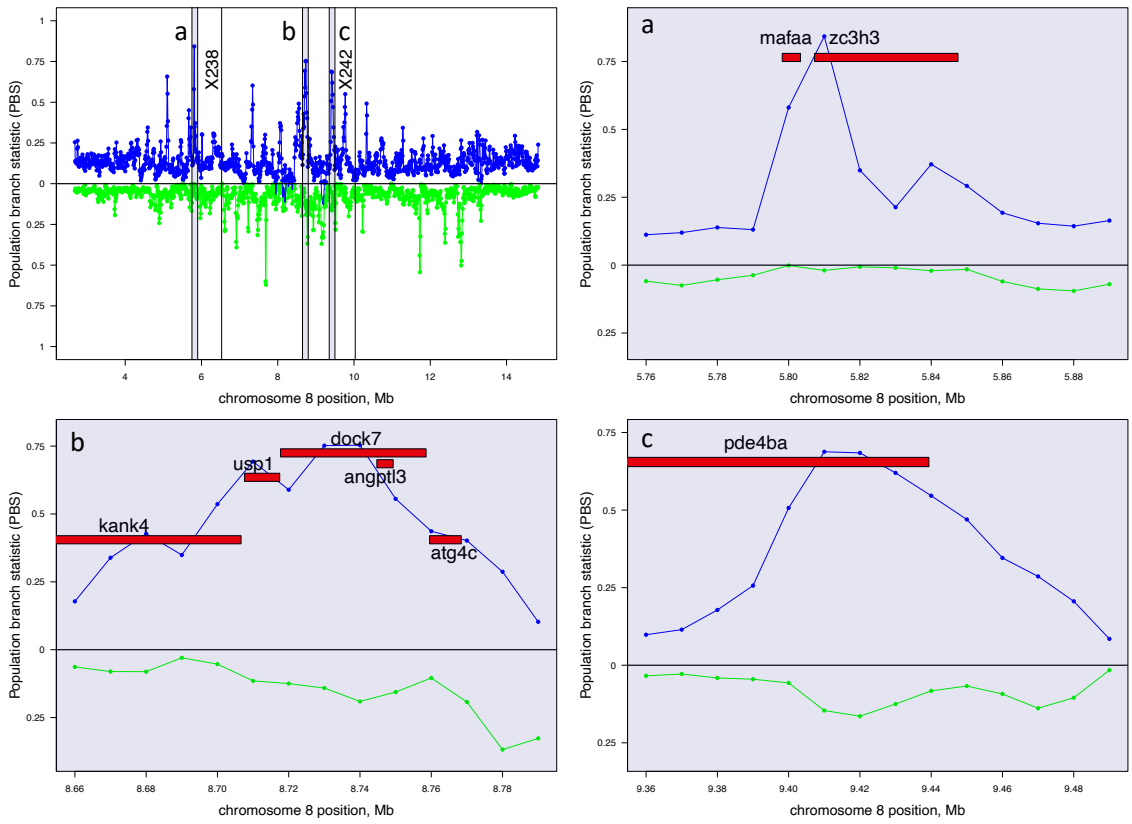

Figure D.11: Strong genomic targets of selection with high population branch statistic distribute around the outstanding snQTLs in Chr 8. Values above the medial line represent higher PBS in Gosling Lake (blue); values below the line represent higher PBS in Roberts Lake (green). Protein-coding genes lie in regions adjacent to three PBS outliers (a, b, c) around snQTLs (markers X238 and X242).

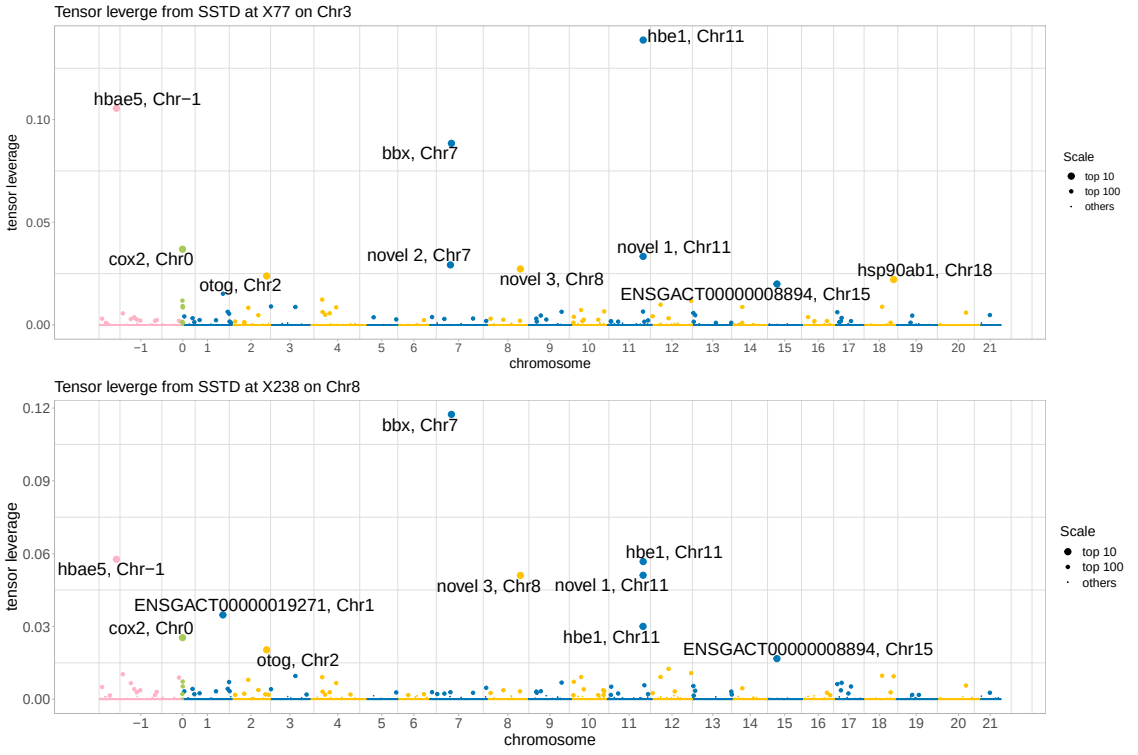

Figure D.12: Joint differential network analysis at snQTLs X77 on Chr 3 (top) and X238 on Chr 8 (bottom) with leverage scores for 10000 genes. Primary genes with top 10 leverage are highlighted with transcription IDs. Mitochondrial genome (MT) and scaffold region are coded as Chr 0 and Chr -1, respectively.

eQTL analysis at snQTL X419

We perform the eQTL analysis at our most significant snQTL, X419 on Chr 18. We regress the residual expression for the top 10,000 genes (used for snQTL main text analysis) onto the genotype of X419. The eQTL results are shown in the Manhattan plot Figure D.14. After multiple testing Bonferroni correction with critical value $0.05/10,000$, only 10 genes (marked on the plot) have expressions significantly affected by X419 genotype. We can conclude that X419 is a cis-eQTL but not a tran-eQTL.

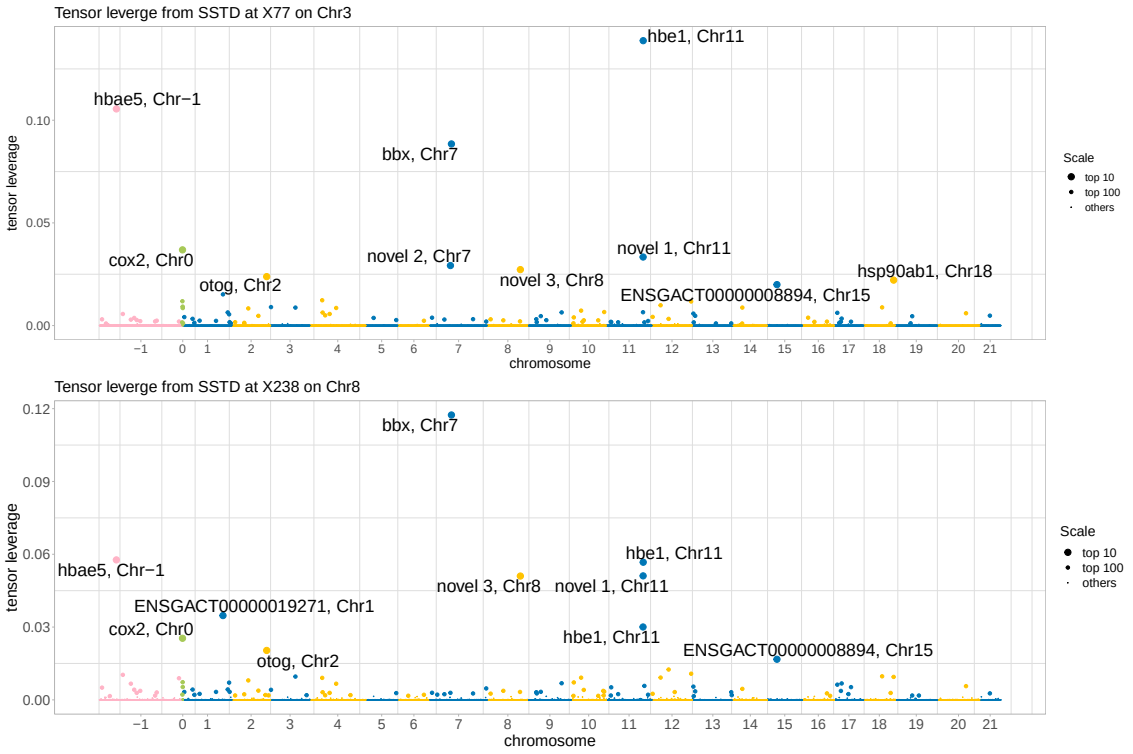

Figure D.13: Joint differential networks at snQTLs (A) X77 on Chr 3 and (B) X238 on Chr 8 with top 100 genes identified in the analysis at X419 on Chr 18. The edge width indicates the connection strength between two genes; the diameter of node indicates the leverage of the genes; the color indicates the enhancement (red) or reduction (blue) of the connection compared with average level. Top 10% strongly connected edges are selected. novel 1: ENSGACT00000018413; novel 2: ENSGACT00000026589; novel 3: ENSGACT00000017116.

Further, we test for and found a marginally significant ($p = 0.051$) association between *ccn6* expression and X419 genotype within the cross type. Figure D.15 indicates the genotype effects to the averaged *ccn6* expression levels.

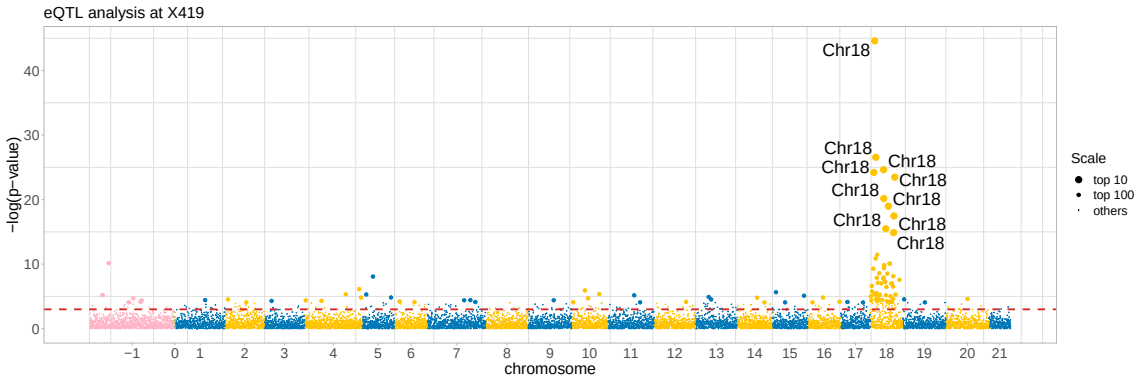

Figure D.14: eQTL analysis at snQTL X419. The red dashed line refers to the critical threshold p-value= 0.05. Top 10 genes with smallest p-values are highlighted with the location text.

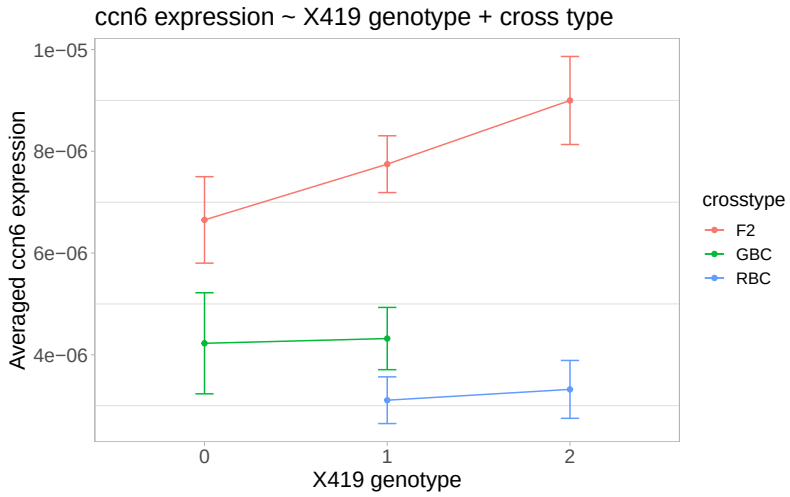

Figure D.15: Averaged *ccn6* expressions under different X419 genotypes and cross types.

Stickleback analysis with different sparsity parameters

We also apply snQTL analysis with different sparsity parameter R to the stickleback data. First, we perform BIC selection for all gene 234 markers with candidate $c \in \{0.2, 0.3, 0.5, 0.7\}$ where $R = c^2 p$. Among 234 markers, 15 markers select $c = 0.2$, 51 markers choose $c = 0.3$, 82 markers choose $c = 0.5$ and 86 markers choose $c = 0.7$. Around 2/3 of markers fit the sparsity assumption in which only a few proportion of genes contribute to the joint network difference. This observation also indicates the necessity to introduce sparsity in the low-rank approximation. Second, we re-run the testing for the key snQTLs with different c's. Table D.5 indicates that our snQTL detection results for stickleback data are robust to the sparsity parameters. Further, we check the leverage and joint differential network estimation at X419 with different c's. The list of top genes with highest leverages and the joint differential network remain the same as the main text results with $c = 0.3$, except slight shifts in leverage values. Hence, we conclude that overall snQTL analyses for stickleback data are robust.

c	0.2	0.3	0.5	0.7
X77	0.008	0.010	0.010	0.010
X81	0.012	0.012	0.012	0.012
X419	0.002	0.008	0.008	0.008
X423	0.008	0.008	0.008	0.008
X238	0.020	0.018	0.018	0.018
X245	0.018	0.020	0.020	0.020

Table D.5: Empirical p-values of top snQTLs obtained by snQTL testing with different sparsity levels. We use $c = 0.3$ for the main text results.

D.3 P-value calibration

In this section, we perform the p-value calibration for stickleback snQTL analysis. The QQ plot (Figure D.16) reveals p-value inflation, suggesting a potential violation

of the exchangeability assumption in the permutation test, likely due to population structure or hidden confounders. The QQ plot Figure D.16, indicates a p-value inflation issue in current results. To calibrate the p-values, we apply genomic control and data whitening transformations.

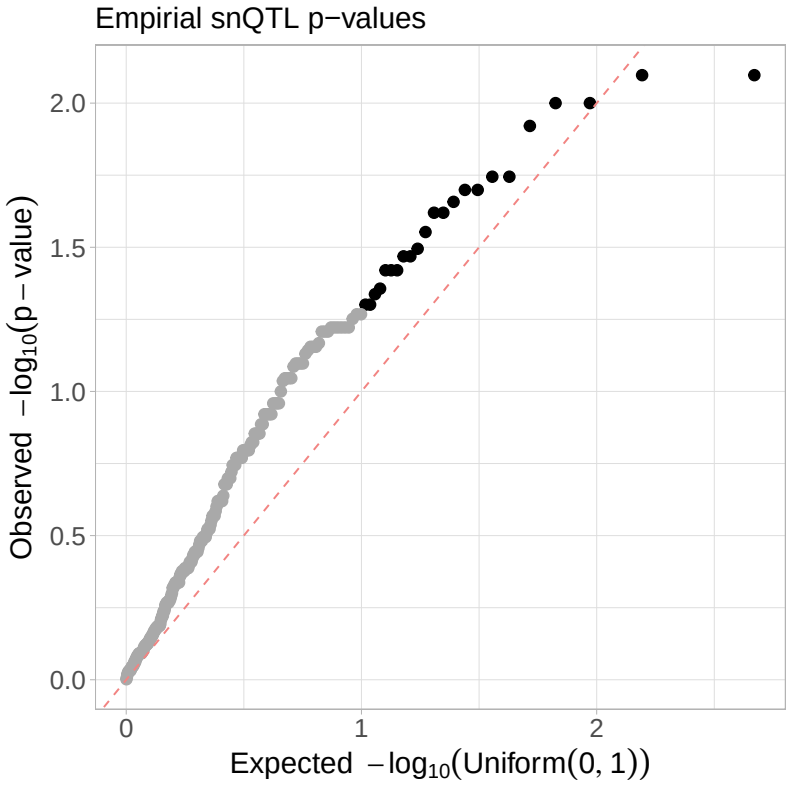

Figure D.16: QQ plot for empirical p-values obtained in stickleback snQTL analysis. P-values smaller than 0.05 are color coded as black, otherwise gray.

Genomic control

We firstly apply classical genomic control (GC, Devlin and Roeder (1999)) to calibrate current empirical p-values. In genetic association studies, GC aims to control the confounding effects of population stratification. In our context, GC is performed as follows:

1. Transform each empirical p-value to a Wald statistic:

$$W_i = (\chi_1^2)^{-1}(1 - pval_i), \quad i = 1, \ldots, 234,$$

where $(\chi_1^2)^{-1}$ refers to the inverse CDF of χ_1^2 distribution, and $pval_i$'s are the empirical p-values of 234 markers.

2. Compute the GC inflation factor:

$$\lambda^{GC} = \frac{\text{median}(W_1, \ldots, W_{234})}{\text{median}(\chi_1^2)}.$$

3. Adjust p-values accordingly:

$$pval_i^{GC} = 1 - \chi_1^2(W_i/\lambda_{GC}).$$

In our stickleback analysis, the QQ plot of GC-adjusted p-values (Figure D.17) reveals moderate deflation, indicating that GC may over-correct for population structure and produce overly conservative results.

Whitening transformation

Unlike GC, the whitening transformation adjusts raw expression counts to decorrelate samples directly. We implement the following pre-processing steps:

(i) Normalization. As in the original pre-processing, raw transcript counts are normalized to account for sequencing depth:

$$X_{ji}^N = 1000000 \times \frac{X_{ji}^0}{\sum_{i=1}^p X_{ji}^0}, \quad \text{for sample } j = 1, \ldots, n,$$

where the scaling factor 1000000 scales up the normalized counts and ensures numerical stability during computation.

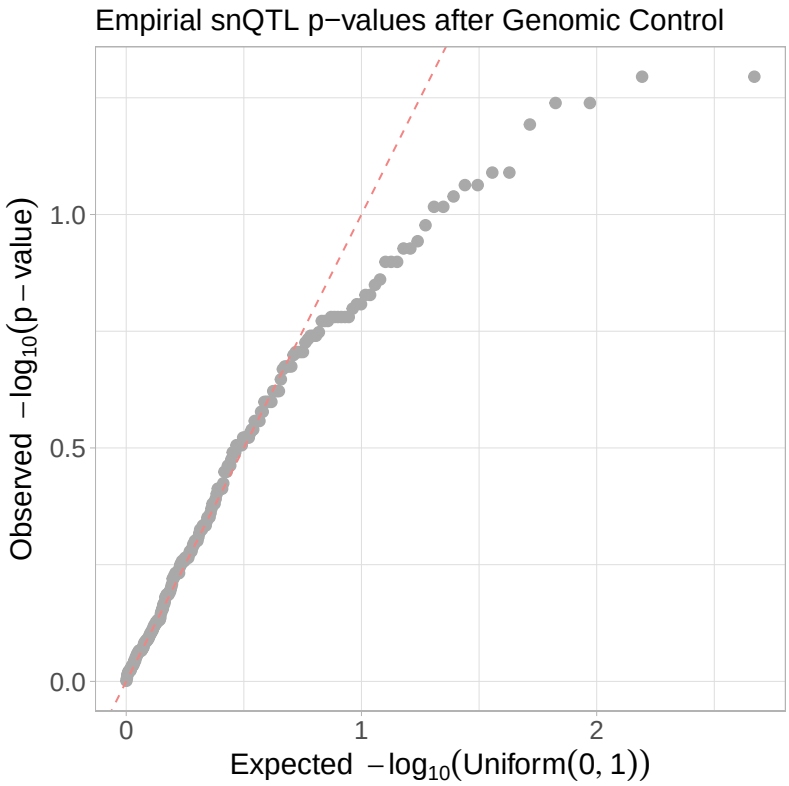

Figure D.17: QQ plot for empirical p-values obtained in stickleback snQTL analysis after genomic control. P-values smaller than 0.05 are color coded as black, otherwise gray.

(ii) Covariance matrix estimation. We represent the dependence structure among individuals via the sample covariance matrix estimation.

$$\Sigma = X^N (X^N)^T.$$

(iii) Generalized linear regression. Accounting the effects of sample dependence and covariate of sex, we build generalized linear regression models:

$$X^N_{:i} = S\beta_i + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}_n(\mathbf{0}, \Sigma), \quad i = 1, \ldots, p,$$

where $X_{:i}^N \in \mathbb{R}^n$ refers to the counts of gene i, $S \in \mathbb{R}^{n\times 3}$ refers to the design matrix for the intercept and 3-level sex factor, $\beta_i \in \mathbb{R}^3$ is the regression coefficient for gene i, and ϵ_i's follow zero-mean multivariate normal distribution with covariance matrix Σ. To remove the sample dependence and covariate effect, we take residuals of the linear regression model with decorrelated response and predictor:

$$X_{:i}^W = \text{residual}(\Sigma^{-1/2}X_{:i}^N \sim \Sigma^{-1/2}S), \text{ equivalently } X_{:i}^W = \Sigma^{-1/2}(X_{:i}^N - S\tilde{\beta}_j),$$

where $\tilde{\beta}_j$ is the least square estimate of the generalized linear regression. Notice that the formula of $X_{:i}^W$ agrees with the general whitening transformation procedure. We call X^W the whitened data.

We re-run the snQTL testing on the whitened stickleback data using 100 permutations per marker, keeping all other settings unchanged. Figure D.18 shows a significant inflation in the tail of the QQ plot while most points align with the diagonal line. This tail inflation pattern confirms the validness of permutation tests and the liability of the significant discoveries. Further, the significant snQTLs with whitened data cluster around Chr 11, Chr 14, and Chr 15 (Figure D.19). In contrast to main text results, these snQTLs locate closely to the eQTL regions identified in the previous study Weber et al. (2022). At this point, we have shown the efficacy of whitening transformation to address the p-value inflation, and we leave further detailed investigations of sample dependence effects on co-expression evolution for future work.

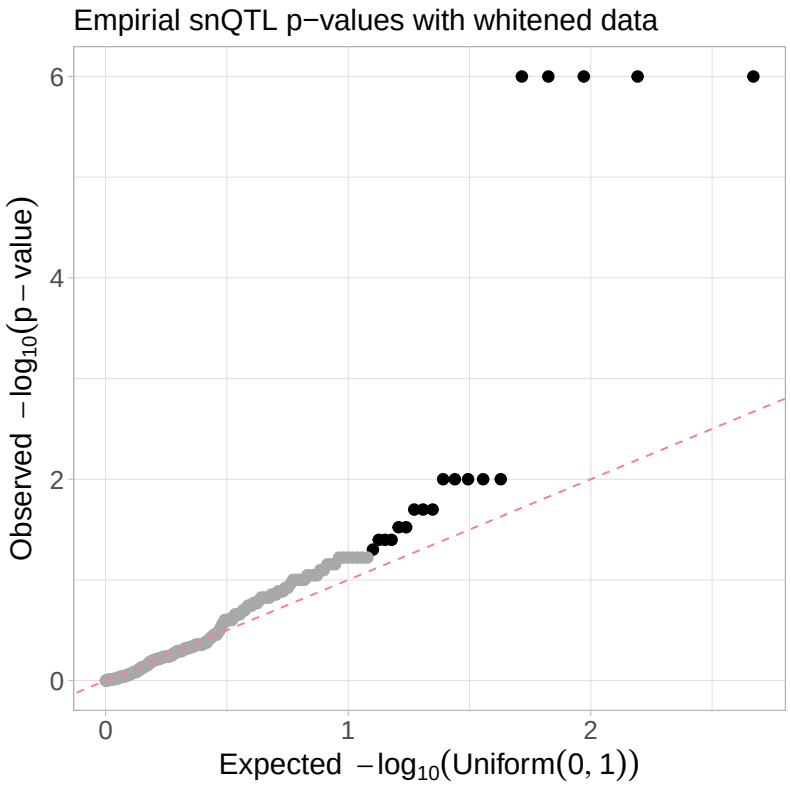

Figure D.18: QQ plot for empirical p-values obtained in stickleback snQTL analysis with whitened data. P-values smaller than 0.05 are color coded as black, otherwise gray.

Figure D.19: Manhattan plot for snQTL testing with tensor statistics and whitened data marks stickleback snQTLs (above pink dashed line with p-values smaller than 0.05), mainly clustered in Chr11, Chr14, and Chr15.

Abbe, Emmanuel. 2017. Community detection and stochastic block models: recent developments. *The Journal of Machine Learning Research* 18(1):6446–6531.

———. 2018. Community detection and stochastic block models: Recent developments. *Journal of Machine Learning Research* 18(177):1–86.

Abbe, Emmanuel, Jianqing Fan, Kaizheng Wang, and Yiqiao Zhong. 2020. Entrywise eigenvector analysis of random matrices with low expected rank. *The Annals of Statistics* 48(3):1452–1474.

Adragni, Kofi P, and R Dennis Cook. 2009. Sufficient dimension reduction and prediction in regression. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 367(1906):4385–4405.

Ahn, Kwangjun, Kangwook Lee, and Changho Suh. 2018. Hypergraph spectral clustering in the weighted stochastic block model. *IEEE Journal of Selected Topics in Signal Processing* 12(5):959–974.

———. 2019. Community recovery in hypergraphs. *IEEE Transactions on Information Theory* 65(10):6561–6579.

Al-Sharoa, Esraa, Mahmood Al-Khassaweneh, and Selin Aviyente. 2018. Tensor based temporal and multilayer community detection for studying brain dynamics during resting state fmri. *IEEE Transactions on Biomedical Engineering* 66(3):695–709.

Allen, Genevera. 2012. Sparse higher-order principal components analysis. In *Artificial intelligence and statistics*, 27–36. PMLR.

Allen, Genevera I, Logan Grosenick, and Jonathan Taylor. 2014. A generalized least-square matrix decomposition. *Journal of the American Statistical Association* 109(505):145–159.

Amini, Arash A, Aiyou Chen, Peter J Bickel, and Elizaveta Levina. 2013. Pseudo-likelihood methods for community detection in large sparse networks. *The Annals of Statistics*.

Anandkumar, Animashree, Rong Ge, Daniel Hsu, Sham M. Kakade, and Matus Telgarsky. 2014. Tensor decompositions for learning latent variable models. *Journal of Machine Learning Research* 15(80):2773–2832.

Auddy, Arnab, Dong Xia, and Ming Yuan. 2024. Tensor methods in high dimensional data analysis: Opportunities and challenges. *arXiv preprint arXiv:2405.18412*.

Baker, Robert L, Wen Fung Leong, Marcus T Brock, Matthew J Rubin, RJ Cody Markelz, Stephen Welch, Julin N Maloof, and Cynthia Weinig. 2019. Integrating transcriptomic network reconstruction and eqtl analyses reveals mechanistic connections between genomic architecture and brassica rapa development. *PLoS Genetics* 15(9):e1008367.

Banerjee, Arindam, Sugato Basu, and Srujana Merugu. 2007. Multi-way clustering on relation graphs. In *Proceedings of the 2007 siam international conference on data mining*, 145–156. SIAM.

Benson, Austin R, David F Gleich, and Jure Leskovec. 2015. Tensor spectral clustering for partitioning higher-order network structures. In *Proceedings of the 2015 siam international conference on data mining*, 118–126. SIAM.

Berthet, Quentin, and Nicolai Baldin. 2020. Statistical and computational rates in graph logistic regression. *Proceedings of the 23th International Conference on Artificial Intelligence and Statistics* 108:2719–2730.

Bickel, Peter J, and Aiyou Chen. 2009. A nonparametric view of network models and newman–girvan and other modularities. *Proceedings of the National Academy of Sciences* 106(50):21068–21073.

Brennan, Matthew, and Guy Bresler. 2020. Reducibility and statistical-computational gaps from secret leakage. In *Proceedings of thirty third conference on learning theory*, vol. 125, 648–847.

Chang, Jinyuan, Wen Zhou, Wen-Xin Zhou, and Lan Wang. 2017. Comparing large covariance matrices under weak conditions on the dependence structure and its application to gene clustering. *Biometrics* 73(1):31–41.

Chen, Han, Garvesh Raskutti, and Ming Yuan. 2019a. Non-convex projected gradient descent for generalized low-rank tensor regression. *The Journal of Machine Learning Research* 20(1):172–208.

Chen, Yuxin, Jianqing Fan, Cong Ma, and Yuling Yan. 2019b. Inference and uncertainty quantification for noisy matrix completion. *Proceedings of the National Academy of Sciences* 116(46):22931–22937.

Chi, Eric C, Brian J Gaines, Will Wei Sun, Hua Zhou, and Jian Yang. 2020. Provable convex co-clustering of tensors. *Journal of Machine Learning Research* 21(214):1–58.

Chi, Eric C, and Tamara G Kolda. 2012. On tensors, sparsity, and nonnegative factorizations. *SIAM Journal on Matrix Analysis and Applications* 33(4):1272–1299.

Chi, Yuejie, Yue M Lu, and Yuxin Chen. 2019. Nonconvex optimization meets low-rank matrix factorization: An overview. *IEEE Transactions on Signal Processing* 67(20):5239–5269.

Chien, I, Chung-Yi Lin, and I-Hsiang Wang. 2018. Community detection in hypergraphs: Optimal statistical limit and efficient algorithms. In *International conference on artificial intelligence and statistics*, 871–879. PMLR.

Chien, I Eli, Chung-Yi Lin, and I-Hsiang Wang. 2019. On the minimax misclassification ratio of hypergraph community detection. *IEEE Transactions on Information Theory* 65(12):8095–8118.

Chisanga, David, Shivakumar Keerthikumar, Suresh Mathivanan, and Naveen Chilamkurti. 2017. Network tools for the analysis of proteomic data. *Proteome Bioinformatics* 177–197.

Consortium, GTEx. 2020. The gtex consortium atlas of genetic regulatory effects across human tissues. *Science* 369(6509):1318–1330.

Costanzo, Michael, Benjamin VanderSluis, Elizabeth N Koch, Anastasia Baryshnikova, Carles Pons, Guihong Tan, Wen Wang, Matej Usaj, Julia Hanchard, Susan D Lee, et al. 2016. A global genetic interaction network maps a wiring diagram of cellular function. *Science* 353(6306):aaf1420.

Crow, Megan, Hamsini Suresh, John Lee, and Jesse Gillis. 2022. Coexpression reveals conserved gene programs that co-vary with cell type across kingdoms. *Nucleic Acids Research* 50(8):4302–4314.

De Lathauwer, Lieven, Bart De Moor, and Joos Vandewalle. 2000. A multilinear singular value decomposition. *SIAM Journal on Matrix Analysis and Applications* 21(4):1253–1278.

Deng, Jiayi, Danyang Huang, Yi Ding, Yingqiu Zhu, Bingyi Jing, and Bo Zhang. 2024. Subsampling spectral clustering for stochastic block models in large-scale networks. *Computational Statistics & Data Analysis* 189:107835.

Desikan, Rahul S, Florent Ségonne, Bruce Fischl, Brian T Quinn, Bradford C Dickerson, Deborah Blacker, Randy L Buckner, Anders M Dale, R Paul Maguire, Bradley T Hyman, Marilyn S Albert, and Ronald J Killiany. 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31:968–980.

Devlin, Bernie, and Kathryn Roeder. 1999. Genomic control for association studies. *Biometrics* 55(4):997–1004.

Ding, Jun, Johann E Gudjonsson, Liming Liang, Philip E Stuart, Yun Li, Wei Chen, Michael Weichenthal, Eva Ellinghaus, Andre Franke, William Cookson, et al. 2010.

Gene expression in skin and lymphoblastoid cells: Refined statistical method reveals extensive overlap in cis-eqtl signals. *The American Journal of Human Genetics* 87(6):779–789.

Dutta, Diptavo, Yuan He, Ashis Saha, Marios Arvanitis, Alexis Battle, and Nilanjan Chatterjee. 2022. Aggregative trans-eqtl analysis detects trait-specific target gene sets in whole blood. *Nature communications* 13(1):4323.

Efron, Bradley, and Robert J Tibshirani. 1994. *An introduction to the bootstrap*. CRC Monographs on Statistics and Applied Probability Series, Chapman and Hall.

Fagerberg, Linn, Björn M Hallström, Per Oksvold, Caroline Kampf, Dijana Djureinovic, Jacob Odeberg, Masato Habuka, Simin Tahmasebpoor, Angelika Danielsson, Karolina Edlund, et al. 2014. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Molecular & cellular proteomics* 13(2):397–406.

Farhadian, Mohammad, Seyed Abbas Rafat, Bahman Panahi, and Christopher Mayack. 2021. Weighted gene co-expression network analysis identifies modules and functionally enriched pathways in the lactation process. *Scientific Reports* 11(1):2367.

Farias, Vivek F, and Andrew A Li. 2019. Learning preferences with side information. *Management Science* 65(7):3131–3149.

Florescu, Laura, and Will Perkins. 2016. Spectral thresholds in the bipartite stochastic block model. In *Proceedings of twenty ninth conference on learning theory*, vol. 49, 943–959.

Fuess, Lauren E, Jesse N Weber, Stijn den Haan, Natalie C Steinel, Kum Chuan Shim, and Daniel I Bolnick. 2021. Between-population differences in constitutive and infection-induced gene expression in threespine stickleback. *Molecular ecology* 30(24):6791–6805.

Gabriel, K Ruben. 1998. Generalised bilinear regression. *Biometrika* 85(3):689–700.

Gahrooei, Mostafa Reisi, Hao Yan, Kamran Paynabar, and Jianjun Shi. 2020. Multiple tensor-on-tensor regression: An approach for modeling processes with heterogeneous sources of data. *Technometrics* 63(2):1–23.

Gaiteri, Chris, Ying Ding, Beverly French, George C Tseng, and Etienne Sibille. 2014. Beyond modules and hubs: the potential of gene coexpression networks for investigating molecular mechanisms of complex brain disorders. *Genes, brain and behavior* 13(1):13–24.

Gao, Chao, Zongming Ma, Anderson Y Zhang, and Harrison H Zhou. 2018. Community detection in degree-corrected block models. *The Annals of Statistics* 46(5): 2153–2185.

Gao, Chao, and Anderson Y Zhang. 2022. Iterative algorithm for discrete structure recovery. *The Annals of Statistics* 50(2):1066–1094.

Ghoshdastidar, Debarghya, and Ambedkar Dukkipati. 2015. A provable generalized tensor spectral method for uniform hypergraph partitioning. In *International conference on machine learning*, 400–409. PMLR.

———. 2017a. Consistency of spectral hypergraph partitioning under planted partition model. *The Annals of Statistics* 45(1):289 – 315.

———. 2017b. Uniform hypergraph partitioning: Provable tensor methods and sampling techniques. *Journal of Machine Learning Research* 18(50):1–41.

Goldfarb, Donald, and Zhiwei Qin. 2014. Robust low-rank tensor recovery: Models and algorithms. *SIAM Journal on Matrix Analysis and Applications* 35(1):225–253.

Greene, Evan, and Jon A Wellner. 2017. Exponential bounds for the hypergeometric distribution. *Bernoulli: official journal of the Bernoulli Society for Mathematical Statistics and Probability* 23(3):1911.

Guo, Xingyi, Weiqiang Lin, Jiandong Bao, Qiuyin Cai, Xiao Pan, Mengqiu Bai, Yuan Yuan, Jiajun Shi, Yaqiong Sun, Mi-Ryung Han, et al. 2018. A comprehensive

cis-eqtl analysis revealed target genes in breast cancer susceptibility loci identified in genome-wide association studies. *The American Journal of Human Genetics* 102(5): 890–903.

Han, Rungang, Yuetian Luo, Miaoyan Wang, and Anru R Zhang. 2022a. Exact clustering in tensor clock model: Statistical optimality and computational limit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 84(5):1666–1698.

Han, Rungang, Rebecca Willett, and Anru Zhang. 2020. An optimal statistical and computational framework for generalized tensor estimation. *The Annals of Statistics, In press. arXiv preprint arXiv:2002.11255*.

Han, Rungang, Rebecca Willett, and Anru R Zhang. 2022b. An optimal statistical and computational framework for generalized tensor estimation. *The Annals of Statistics* 50(1):1–29.

Hao, Botao, Boxiang Wang, Pengyuan Wang, Jingfei Zhang, Jian Yang, and Will Wei Sun. 2021. Sparse tensor additive regression. *Journal of Machine Learning Research* 22(64):1–43.

Hao, Botao, Anru Zhang, and Guang Cheng. 2020. Sparse and low-rank tensor estimation via cubic sketchings. *IEEE Transactions on Information Theory* 66(9): 5927–5964.

Hitchcock, Frank L. 1927. The expression of a tensor or a polyadic as a sum of products. *Journal of Mathematics and Physics* 6(1-4):164–189.

Hoff, Peter D. 2005. Bilinear mixed-effects models for dyadic data. *Journal of the American Statistical Association* 100(469):286–295.

Hong, David, Tamara G Kolda, and Jed A Duersch. 2020. Generalized canonical polyadic tensor decomposition. *SIAM Review* 62(1):133–163.

Hore, Victoria, Ana Viñuela, Alfonso Buil, Julian Knight, Mark I McCarthy, Kerrin Small, and Jonathan Marchini. 2016. Tensor decomposition for multiple-tissue gene expression experiments. *Nature genetics* 48(9):1094.

Hoskins, Jason W, Charles C Chung, Aidan O'Brien, Jun Zhong, Katelyn Connelly, Irene Collins, Jianxin Shi, and Laufey T Amundadottir. 2021. Inferred expression regulator activities suggest genes mediating cardiometabolic genetic signals. *PLOS Computational Biology* 17(11):e1009563.

Hu, Jiaxin. 2025. Large-scale multiway clustering with seeded clustering. In *The second conference on parsimony and learning (proceedings track)*.

Hu, Jiaxin, Chanwoo Lee, and Miaoyan Wang. 2022. Generalized tensor decomposition with features on multiple modes. *Journal of Computational and Graphical Statistics* 31(1):204–218.

Hu, Jiaxin, and Miaoyan Wang. 2023. Multiway spherical clustering via degree-corrected tensor block models. *IEEE Transactions on Information Theory* 69(6).

Hu, Jiaxin, Jesse N Weber, Lauren E Fuess, Natalie C Steinel, Daniel I Bolnick, and Miaoyan Wang. 2024. A spectral framework to map qtls affecting joint differential networks of gene co-expression. *bioRxiv*.

Ingalhalikar, Madhura, Alex Smith, Drew Parker, Theodore D Satterthwaite, Mark A Elliott, Kosha Ruparel, Hakon Hakonarson, Raquel E Gur, Ruben C Gur, and Ragini Verma. 2014. Sex differences in the structural connectome of the human brain. *Proceedings of the National Academy of Sciences* 111(2):823–828.

Kang, Mingon, Chunling Zhang, Hyung-Wook Chun, Chris Ding, Chunyu Liu, and Jean Gao. 2015. eqtl epistasis: detecting epistatic effects and inferring hierarchical relationships of genes in biological pathways. *Bioinformatics* 31(5): 656–664.

Kazemi, Ehsan, S Hamed Hassani, and Matthias Grossglauser. 2015. Growing a graph matching from a handful of seeds. *Proceedings of the VLDB Endowment* 8(10):1010–1021.

Ke, Zheng Tracy, Feng Shi, and Dong Xia. 2019. Community detection for hypergraph networks via regularized tensor power iteration. *arXiv preprint arXiv:1909.06503*.

Kim, Chiheon, Afonso S Bandeira, and Michel X Goemans. 2018. Stochastic block model for hypergraphs: Statistical limits and a semidefinite programming approach. *arXiv preprint arXiv:1807.02884*.

Kolberg, Liis, Nurlan Kerimov, Hedi Peterson, and Kaur Alasoo. 2020. Co-expression analysis reveals interpretable gene modules controlled by trans-acting genetic variants. *Elife* 9:e58705.

Kolda, Tamara G, and Brett W Bader. 2009. Tensor decompositions and applications. *SIAM Review* 51(3):455–500.

Koniusz, Piotr, and Anoop Cherian. 2016. Sparse coding for third-order super-symmetric tensor descriptors with application to texture recognition. In *Proceedings of the ieee conference on computer vision and pattern recognition*, 5395–5403.

Lang, Joel, and James Henderson. 2013. Graph-based seed set expansion for relation extraction using random walk hitting times. In *Proceedings of the 2013 conference of the north american chapter of the association for computational linguistics: Human language technologies*, 772–776.

Lee, Chanwoo, and Miaoyan Wang. 2021. Beyond the signs: Nonparametric tensor completion via sign series. *arXiv preprint arXiv:2102.00384*.

———. 2024. Statistical and computational efficiency for smooth tensor estimation with unknown permutations. *Journal of the American Statistical Association* 1–14.

Lee, Sang Hoon, José Manuel Magallanes, and Mason A Porter. 2017. Time-dependent community structure in legislation cosponsorship networks in the congress of the republic of peru. *Journal of Complex Networks* 5(1):127–144.

Li, Gen. 2020. Generalized co-clustering analysis via regularized alternating least squares. *Computational Statistics & Data Analysis* 150:106989.

Li, Jun, and Song Xi Chen. 2012. Two sample tests for high-dimensional covariance matrices. *The Annals of Statistics* 40(2):908 – 940.

Li, Lexin, and Xin Zhang. 2017. Parsimonious tensor response regression. *Journal of the American Statistical Association* 112(519):1131–1146.

Li, Shuang, Katharina T Schmid, Dylan H de Vries, Maryna Korshevniuk, Corinna Losert, Roy Oelen, Irene V van Blokland, sc-eQTLgen Consortium BIOS Consortium, Hilde E Groot, Morris A Swertz, et al. 2023. Identification of genetic variants that impact gene co-expression relationships using large-scale single-cell data. *Genome Biology* 24(1):80.

Lock, Eric F. 2018. Tensor-on-tensor regression. *Journal of Computational and Graphical Statistics* 27(3):638–647.

Lock, Eric F, Katherine A Hoadley, James Stephen Marron, and Andrew B Nobel. 2013. Joint and individual variation explained (JIVE) for integrated analysis of multiple data types. *The Annals of Applied Statistics* 7(1):523.

Lock, Eric F, and Gen Li. 2018. Supervised multiway factorization. *Electronic Journal of Statistics* 12(1):1150.

Lohman, Brian K, Natalie C Steinel, Jesse N Weber, and Daniel I Bolnick. 2017. Gene expression contributes to the recent evolution of host resistance in a model host parasite system. *Frontiers in immunology* 1071.

Lu, Yu, and Harrison H Zhou. 2016. Statistical and computational guarantees of lloyd's algorithm and its variants. *arXiv preprint arXiv:1612.02099*.

Luo, Chongliang, Jian Liang, Gen Li, Fei Wang, Changshui Zhang, Dipak K Dey, and Kun Chen. 2018. Leveraging mixed and incomplete outcomes via reduced-rank modeling. *Journal of Multivariate Analysis* 167:378–394.

Luo, Yuetian, and Anru R Zhang. 2021. Low-rank tensor estimation via riemannian gauss-newton: Statistical optimality and second-order convergence. *arXiv preprint arXiv:2104.12031*.

———. 2022. Tensor clustering with planted structures: Statistical optimality and computational limits. *The Annals of Statistics* 50(1):584–613.

Majewski, Jacek, and Tomi Pastinen. 2011. The study of eqtl variations by rna-seq: from snps to phenotypes. *Trends in Genetics* 27(2):72–79.

McCullagh, P., and J.A. Nelder. 1989. *Generalized linear models, second edition*. CRC Monographs on Statistics and Applied Probability Series, Chapman and Hall.

Meilă, Marina. 2012. Local equivalences of distances between clusterings—a geometric perspective. *Machine Learning* 86(3):369–389.

Miller, Dustyn S, and Malini Sen. 2007. Potential role of wisp3 (ccn6) in regulating the accumulation of reactive oxygen species. *Biochemical and biophysical research communications* 355(1):156–161.

Mu, Cun, Bo Huang, John Wright, and Donald Goldfarb. 2014. Square deal: Lower bounds and improved relaxations for tensor recovery. In *Proceedings of the 31st international conference on machine learning*, vol. 32, 73–81.

Mukherjee, Soumendu Sundar, Purnamrita Sarkar, and Peter J Bickel. 2021. Two provably consistent divide-and-conquer clustering algorithms for large networks. *Proceedings of the National Academy of Sciences* 118(44):e2100482118.

Nica, Alexandra C, and Emmanouil T Dermitzakis. 2013. Expression quantitative trait loci: present and future. *Philosophical Transactions of the Royal Society B: Biological Sciences* 368(1620):20120362.

Nickel, Maximilian, Volker Tresp, and Hans-Peter Kriegel. 2011. A three-way model for collective learning on multi-relational data. *Proceedings of the 28th International Conference on Machine Learning* 809–816.

Oldham, Michael C, Steve Horvath, and Daniel H Geschwind. 2006. Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *Proceedings of the National Academy of Sciences* 103(47):17973–17978.

Ovens, Katie, B Frank Eames, and Ian McQuillan. 2021. Comparative analyses of gene co-expression networks: Implementations and applications in the study of evolution. *Frontiers in Genetics* 12:695399.

Padhan, Deepesh Kumar, Archya Sengupta, Milan Patra, Ananya Ganguly, Sushil Kumar Mahata, and Malini Sen. 2020. Ccn6 regulates mitochondrial respiratory complex assembly and activity. *The FASEB Journal* 34(9):12163–12176.

Pantelis, Christos, George N Papadimitriou, Sergi Papiol, Elena Parkhomenko, Michele T Pato, Tiina Paunio, Milica Pejovic-Milovancevic, Diana O Perkins, Olli Pietiläinen, et al. 2014. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511(7510):421–427.

Park, Kwangmoon, and Sündüz Keleş. 2024. Joint tensor modeling of single cell 3d genome and epigenetic data with muscle. *Journal of the American Statistical Association* 119(548):2464–2477.

Perbal, Bernard. 2018. The concept of the ccn protein family revisited: a centralized coordination network. *Journal of cell communication and signaling* 12:3–12.

Peters, James E, Paul A Lyons, James C Lee, Arianne C Richard, Mary D Fortune, Paul J Newcombe, Sylvia Richardson, and Kenneth GC Smith. 2016. Insight into genotype-phenotype associations through eqtl mapping in multiple cell types in health and immune-mediated disease. *PLoS genetics* 12(3):e1005908.

Rabusseau, Guillaume, and Hachem Kadri. 2016. Low-rank regression with tensor responses. *Proceedings of the 30th International Conference on Neural Information Processing Systems* 29:1875–1883.

Rahmani, Mostafa, Andre Beckus, Adel Karimian, and George K Atia. 2020. Scalable and robust community detection with randomized sketching. *IEEE Transactions on Signal Processing* 68:962–977.

Raskutti, Garvesh, Ming Yuan, and Han Chen. 2019. Convex regularization for high-dimensional multiresponse tensor regression. *The Annals of Statistics* 47(3): 1554–1584.

Replication, DIAbetes Genetics, Meta analysis (DIAGRAM) Consortium, Asian Genetic Epidemiology Network Type 2 Diabetes (AGEN-T2D) Consortium, South Asian Type 2 Diabetes (SAT2D) Consortium, Mexican American Type 2 Diabetes (MAT2D) Consortium, Type 2 Diabetes Genetic Exploration by Next-generation sequencing in multi Ethnic Samples (T2D-GENES) Consortium, Anubha Mahajan, Min Jin Go, Weihua Zhang, Jennifer E Below, Kyle J Gaulton, et al. 2014. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nature genetics* 46(3):234–244.

Repudi, Srinivasa Rao, Milan Patra, and Malini Sen. 2013. Wisp3–igf1 interaction regulates chondrocyte hypertrophy. *Journal of cell science* 126(7):1650–1658.

Rigollet, Phillippe, and Jan-Christian Hütter. 2015. High dimensional statistics. *Lecture notes for course 18S997*.

Rudelson, Mark, and Roman Vershynin. 2010. Non-asymptotic theory of random matrices: extreme singular values. In *Proceedings of the international congress of mathematicians 2010 (icm 2010) (in 4 volumes) vol. i: Plenary lectures and ceremonies vols. ii–iv: Invited lectures*, 1576–1602. World Scientific.

Ruprecht, Colin, Neha Vaid, Sebastian Proost, Staffan Persson, and Marek Mutwil. 2017. Beyond genomics: studying evolution with gene coexpression networks. *Trends in plant science* 22(4):298–307.

Sayan Chakrabarty, Srijan Sengupta, and Yuguo Chen. 2023. Subsampling based community detection for large networks. *Statistica Sinica*.

Smith, Shaden, Jee W. Choi, Jiajia Li, Richard Vuduc, Jongsoo Park, Xing Liu, and George Karypis. 2017. FROSTT: The formidable repository of open sparse tensors and tools.

Song, Qingquan, Hancheng Ge, James Caverlee, and Xia Hu. 2019. Tensor completion algorithms in big data analytics. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 13(1):1–48.

Song, Yiran, Chenyang Li, Yuxin Luo, Jinbo Guo, Yaxing Kang, Fengrong Yin, Lihong Ye, Donglei Sun, Jun Yu, and Xiaolan Zhang. 2023. Ccn6 improves hepatic steatosis, inflammation, and fibrosis in non-alcoholic steatohepatitis. *Liver International* 43(2):357–369.

Srivastava, Muni S, Tatjana von Rosen, and Dietrich Von Rosen. 2008. Models with a kronecker product covariance structure: Estimation and testing. *Mathematical Methods of Statistics* 17(4):357–370.

Stuart, Joshua M, Eran Segal, Daphne Koller, and Stuart K Kim. 2003. A gene-coexpression network for global discovery of conserved genetic modules. *science* 302(5643):249–255.

Su, Yansen, Bangju Wang, and Xingyi Zhang. 2017. A seed-expanding method based on random walks for community detection in networks with ambiguous community structures. *Scientific reports* 7(1):41830.

Sun, Will Wei, and Lexin Li. 2017. STORE: sparse tensor response regression and neuroimaging analysis. *The Journal of Machine Learning Research* 18(1):4908–4944.

Sun, Will Wei, Junwei Lu, Han Liu, and Guang Cheng. 2017. Provable sparse tensor decomposition. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 79(3):899–916.

Swanson-Wagner, Ruth A, Rhonda DeCook, Yi Jia, Tim Bancroft, Tieming Ji, Xuefeng Zhao, Dan Nettleton, and Patrick S Schnable. 2009. Paternal dominance of trans-eqtl influences gene expression patterns in maize hybrids. *science* 326(5956): 1118–1120.

Tarzanagh, Davoud Ataee, and George Michailidis. 2019. Regularized and smooth double core tensor factorization for heterogeneous data. *arXiv preprint arXiv:1911.10454*.

Traag, Vincent A, and Lovro Šubelj. 2023. Large network community detection by fast label propagation. *Scientific Reports* 13(1):2701.

Tucker, Ledyard R. 1966. Some mathematical notes on three-mode factor analysis. *Psychometrika* 31(3):279–311.

Van Der Wijst, Monique GP, Harm Brugge, Dylan H De Vries, Patrick Deelen, Morris A Swertz, LifeLines Cohort Study, BIOS Consortium, and Lude Franke. 2018. Single-cell rna sequencing identifies celltype-specific cis-eqtls and co-expression qtls. *Nature genetics* 50(4):493–497.

Van Essen, David C, Stephen M Smith, Deanna M Barch, Timothy EJ Behrens, Essa Yacoub, Kamil Ugurbil, and WU-Minn HCP Consortium. 2013. The WU-Minn human connectome project: An overview. *Neuroimage* 80:62–79.

Visscher, Peter M, Matthew A Brown, Mark I McCarthy, and Jian Yang. 2012. Five years of gwas discovery. *The American Journal of Human Genetics* 90(1):7–24.

Von Luxburg, Ulrike. 2007. A tutorial on spectral clustering. *Statistics and computing* 17:395–416.

Võsa, Urmo, Annique Claringbould, Harm-Jan Westra, Marc Jan Bonder, Patrick Deelen, Biao Zeng, Holger Kirsten, Ashis Saha, Roman Kreuzhuber, Seyhan Yazar, et al. 2021. Large-scale cis-and trans-eqtl analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nature genetics* 53(9):1300–1310.

Wan, Dan, Xu Wang, Qinghua Wu, Pingping Lin, Yuanhu Pan, Adeel Sattar, Lingli Huang, Ijaz Ahmad, Yuanyuan Zhang, and Zonghui Yuan. 2015. Integrated transcriptional and proteomic analysis of growth hormone suppression mediated by trichothecene t-2 toxin in rat gh3 cells. *Toxicological Sciences* 147(2):326–338.

Wang, Jiangzhou, Jingfei Zhang, Binghui Liu, Ji Zhu, and Jianhua Guo. 2023. Fast network community detection with profile-pseudo likelihood methods. *Journal of the American Statistical Association* 118(542):1359–1372.

Wang, Lu, Daniele Durante, Rex E Jung, and David B Dunson. 2017. Bayesian network–response regression. *Bioinformatics* 33(12):1859–1866.

Wang, Miaoyan, Jonathan Fischer, and Yun S Song. 2019. Three-way clustering of multi-tissue multi-individual gene expression data using semi-nonnegative tensor decomposition. *The Annals of Applied Statistics* 13(2):1103–1127.

Wang, Miaoyan, and Lexin Li. 2020a. Learning from binary multiway data: Probabilistic tensor decomposition and its statistical optimality. *Journal of Machine Learning Research* 21(154):1–38.

———. 2020b. Learning from binary multiway data: Probabilistic tensor decomposition and its statistical optimality. *Journal of Machine Learning Research* 21(154): 1–38.

Wang, Miaoyan, and Yun Song. 2017. Tensor decompositions via two-mode higher-order SVD (HOSVD). In *Artificial intelligence and statistics*, 614–622.

Wang, Miaoyan, and Yuchen Zeng. 2019. Multiway clustering via tensor block models. In *Advances in neural information processing systems*, vol. 32.

Wang, Xinchen, and David B Goldstein. 2020. Enhancer domains predict gene pathogenicity and inform gene discovery in complex disease. *The American Journal of Human Genetics* 106(2):215–233.

Wang, Xuran, David Choi, and Kathryn Roeder. 2021. Constructing local cell-specific networks from single-cell data. *Proceedings of the National Academy of Sciences* 118(51):e2113178118.

Weber, Jesse N, Natalie C Steinel, Foen Peng, Kum Chuan Shim, Brian K Lohman, Lauren E Fuess, Swapna Subramanian, Stephen P De Lisle, and Daniel I Bolnick. 2022. Evolutionary gain and loss of a pathological immune response to parasitism. *Science* 377(6611):1206–1211.

Wei, Jun, Yu Fang, Hao Jiang, Xing-ting Wu, Jing-hong Zuo, Xian-chun Xia, Jin-quan Li, Benjamin Stich, Hong Cao, and Yong-xiu Liu. 2022. Combining qtl mapping and gene co-expression network analysis for prediction of candidate genes and molecular network related to yield in wheat. *BMC Plant Biology* 22(1): 1–14.

Westra, Harm-Jan, Marjolein J Peters, Tõnu Esko, Hanieh Yaghootkar, Claudia Schurmann, Johannes Kettunen, Mark W Christiansen, Benjamin P Fairfax, Katharina Schramm, Joseph E Powell, et al. 2013. Systematic identification of trans eqtls as putative drivers of known disease associations. *Nature genetics* 45(10): 1238–1243.

Whang, Joyce Jiyoung, David F Gleich, and Inderjit S Dhillon. 2013. Overlapping community detection using seed set expansion. In *Proceedings of the 22nd acm international conference on information & knowledge management*, 2099–2108.

———. 2016. Overlapping community detection using neighborhood-inflated seed expansion. *IEEE Transactions on Knowledge and Data Engineering* 28(5):1272–1284.

Witten, Daniela M, Robert Tibshirani, and Trevor Hastie. 2009. A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. *Biostatistics* 10(3):515–534.

Wu, Shihao, Zhe Li, and Xuening Zhu. 2023. A distributed community detection algorithm for large scale networks under stochastic block models. *Computational Statistics & Data Analysis* 187:107794.

Xia, Dong, and Ming Yuan. 2021. Effective tensor sketching via sparsification. *IEEE Transactions on Information Theory* 67(2):1356–1369.

Xu, Bingying, Zheng Liang, Yan Jia, Bin Zhou, and Yi Han. 2012. Local community detection using seeds expansion. In *2012 second international conference on cloud and green computing*, 557–562. IEEE.

Yao, Douglas W, Luke J O'connor, Alkes L Price, and Alexander Gusev. 2020. Quantifying genetic effects on disease mediated by assayed gene expression levels. *Nature genetics* 52(6):626–633.

Yeger, Herman, and Bernard Perbal. 2016. Ccn family of proteins: critical modulators of the tumor cell microenvironment. *Journal of cell communication and signaling* 10:229–240.

Yu, Liren, Jiaming Xu, and Xiaojun Lin. 2021. Graph matching with partially-correct seeds. *Journal of Machine Learning Research* 22(280):1–54.

Yuan, Kai, Tao Zeng, and Luonan Chen. 2022a. Interpreting functional impact of genetic variations by network qtl for genotype–phenotype association study. *Frontiers in Cell and Developmental Biology* 9:720321.

Yuan, Mingao, Ruiqi Liu, Yang Feng, and Zuofeng Shang. 2022b. Testing community structure for hypergraphs. *The Annals of Statistics* 50(1):147–169.

Yun, Se-Young, and Alexandre Proutiere. 2016. Optimal cluster recovery in the labeled stochastic block model. In *Advances in neural information processing systems*, vol. 29.

Zhang, Anru, and Dong Xia. 2018. Tensor SVD: Statistical and computational limits. *IEEE Transactions on Information Theory* 64(11):7311–7338.

Zhang, Bin, and Steve Horvath. 2005. A general framework for weighted gene co-expression network analysis. *Statistical applications in genetics and molecular biology* 4(1).

Zhang, Hai, Xiao Guo, and Xiangyu Chang. 2022a. Randomized spectral clustering in large-scale stochastic block models. *Journal of Computational and Graphical Statistics* 31(3):887–906.

Zhang, Jingfei, Will Wei Sun, and Lexin Li. 2018. Network response regression for modeling population of networks with covariates. *arXiv preprint arXiv:1810.03192*.

Zhang, Sheng, Rui Song, Wenbin Lu, and Ji Zhu. 2022b. Distributed community detection in large networks. *arXiv preprint arXiv:2203.06509*.

Zhang, Weitong, Ronghua Shang, and Licheng Jiao. 2023. Large-scale community detection based on core node and layer-by-layer label propagation. *Information Sciences* 632:1–18.

Zhang, Zhengwu, Genevera I Allen, Hongtu Zhu, and David Dunson. 2019. Tensor network factorizations: Relationships between brain structural connectomes and traits. *Neuroimage* 197:330–343.

Zhao, He, Guixia Liu, and Xintian Cao. 2023. A seed expansion-based method to identify essential proteins by integrating protein–protein interaction sub-networks and multiple biological characteristics. *BMC bioinformatics* 24(1):452.

Zhao, Wei, Peter Langfelder, Tova Fuller, Jun Dong, Ai Li, and Steve Hovarth. 2010. Weighted gene coexpression network analysis: state of the art. *Journal of biopharmaceutical statistics* 20(2):281–300.

Zhen, Yaoming, and Junhui Wang. 2023. Community detection in general hypergraph via graph embedding. *Journal of the American Statistical Association* 118(543): 1620–1629.

Zhernakova, Daria V, Patrick Deelen, Martijn Vermaat, Maarten Van Iterson, Michiel Van Galen, Wibowo Arindrarto, Peter Van't Hof, Hailiang Mei, Freerk Van Dijk, Harm-Jan Westra, et al. 2017. Identification of context-dependent expression quantitative trait loci in whole blood. *Nature genetics* 49(1):139–145.

Zhou, Hua, Lexin Li, and Hongtu Zhu. 2013. Tensor regression with applications in neuroimaging data analysis. *Journal of the American Statistical Association* 108(502): 540–552.

Zhu, Lingxue, Jing Lei, Bernie Devlin, and Kathryn Roeder. 2017. Testing high-dimensional covariance matrices, with application to detecting schizophrenia risk genes. *The annals of applied statistics* 11(3):1810.

Zhu, Zhihong, Futao Zhang, Han Hu, Andrew Bakshi, Matthew R Robinson, Joseph E Powell, Grant W Montgomery, Michael E Goddard, Naomi R Wray, Peter M Visscher, et al. 2016. Integration of summary data from gwas and eqtl studies predicts complex trait gene targets. *Nature genetics* 48(5):481–487.