2020

Estimation and Clustering in Network and Indirect Data

Ramchandra Rimal
University of Central Florida

ESTIMATION AND CLUSTERING FOR NETWORK AND INDIRECT DATA

by

RAMCHANDRA RIMAL
M.S University of Central Florida, 2017
M.A Tribhuvan University, 2010

A Dissertation submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy
in the Department of Mathematical Sciences
in the College of College of Sciences
at the University of Central Florida
Orlando, Florida

Summer Term
2020

Major Professor: Marianna Pensky

ABSTRACT

The first part of the dissertation studies a density deconvolution problem with small Berkson errors. In this setting, the data is not available directly but rather in the form of convolution and one needs to estimate the convolution of the unknown density with Berkson errors. While it is known that the Berkson errors improve the precision of the reconstruction, it does not necessarily happen when Berkson errors are small. Furthermore, the choice of bandwidth in density estimation has been an open problem so far. In this dissertation, we provide an in-depth study of the choice of the bandwidth which leads to the optimal error rates.

The second part of the dissertation studies a generative network model, the so-called Popularity Adjusted Block Model (PABM) introduced by Sengupta and Chen (2018). The PABM generalizes popular graph generative models such as the Stochastic Block Model (SBM) and the Degree Corrected Block Model (DCBM). The advantages of the PABM is that, unlike mixed membership models or the DCBM, it does not rely on any identifiability conditions, and leads to more flexible spectral properties. We expand the theory of PABM to the case of an arbitrary number of communities which possibly grows with a number of nodes in the network and is not assumed to be known. We produce the estimators of the probability matrix and the community structure and provide non-asymptotic upper bounds for the estimation and the clustering errors.

Majority of real-life networks are sparse, in the sense that they have few high degree nodes while the rest of the nodes have low degrees. Since the SBM and DCBM do not allow to set any probabilities of connections to zero, they model sparsity by enforcing the maximum connection probability to be bounded above by a small quantity which precludes existence of high degree nodes. On the

contrary, the PABM allows modeling some of the probabilities of connections between the nodes as identical zeros while maintaining the rest of the probabilities non-negligible. This leads to the Sparse Popularity Adjusted Block Model (SPABM). The SPABM reduces the size of parameter set and leads to improved precision of estimation and clustering. We produce the estimators of the probability matrix and the community structure in SPABM. Finally, we provide non-asymptotic upper bounds for the estimation and the clustering errors.

To my beloved wife Rojina and my daughter Reeva

.

ACKNOWLEDGMENTS

First of all, I would like to express my gratitude to my dissertation committee members Dr. Marianna Pensky, Dr. Jason Swanson, Dr. Ullas Bagci, and Dr. Teng Zhang. A very special thanks to Dr. Marianna Pensky, my advisor, for her magnanimity, for motivating me, and providing insightful comments on my works with a great amount of patience throughout the whole process. I would not have achieved this level without her wise and sensible guidance.

I extend my thanks to the chair of the mathematics department Dr. Xin Li, the co-chair Dr. Joseph Brennan, and the graduate coordinator Dr. Qiyu Sun. I thank all the faculty members with whom I took courses, for their countless help throughout the period. I appreciate the help from Dr. Marianna Pensky, Dr. Qiyu Sun, Dr. Ullas Bagci, Dr. Dorin Dutkay, Mr. Rachid Ait Maalem Lahcen, and Mr. Keith Carlson during the job search process. I am thankful to the staff members of the mathematics department, namely, Norma Robles, Doreen Goulding, Linda Perez-Rodriguez and Patricia O'Leary for providing help on every possible moment.

I was very fortunate to have friends like Ted Juste, Nathaniel Adu, Christian Bosse, Jordan D'Abruzzo, Blake Wallace, Majid Noroozi, Zingmei Zhang, Arielle Gaudiello, Nazar Emirov, Jason Bentley, and Rasika Rajapakshage. Together with them, I enjoyed the journey towards the Ph.D. in Mathematics. Working together with them made me better in understanding the meaning of mathematics as well as of life.

My deep respect goes towards my father Krishna Prasad Rimal, who persistently provide me with better opportunities beyond his financial ability. He is the best father as well as the best friend I could have. His enormous dedication to help me by any means cannot be praised with words. The intense blessings and love of my mother Radhika Rimal inside my heart provides me with vigorous energy to face the challenges. I always enjoy friendly suggestions from my loving brother

Dharma Raj Rimal. This work could not have been completed without the unconditional love and understanding of my lovely wife. Her strong help, encouragement, and patience during the hardship gave me the strength to keep going.

My casual meeting with professors such as Dr. Ram Mohapatra, Dr. Juhair Nashed, Dr. Piotr Mikusinski, Dr. Kuppalapalle Vajravelu, in the hallway, on the stairs or inside the elevators, always added some useful information towards my career. Finally, I would like to thank all the universities, and organizations that provided me support to attend the conferences, workshops and summer schools during my graduate study.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

In this dissertation, we consider estimation and clustering for two types of data: indirect convolution data and network data.

The first part of the dissertation deals with the deconvolution problem. Deconvolution problems occur in many fields of nonparametric statistics, for example, density estimation based on contaminated data, nonparametric regression with errors-in-variables, image and signal deblurring. Those topics have received more and more attention during the last two decades. The real life applications of deconvolution procedures are in econometrics, astronomy, biometrics, medical statistics, image reconstruction. The general deconvolution problem for density estimation is one where a sample of independent and identically distributed (i.i.d.) variables $Y_1, ..., Y_n$ is observed with random measurement error. The observations are generated by the model

$$Y_j = X_j + \varepsilon_j, \;\; X_j \sim f, \;\; \varepsilon_j \sim g.$$

Here the problem is to estimate the density f of X_j which are unknown, ε_j are called error variable independent of X_j. g is known and is known as error density or blurring density.

In many real life problems one is interested in distribution of a certain variable which can be observed only indirectly. Mathematically, this leads to a density deconvolution problem where one needs to estimate the pdf of a variable X on the basis of observations of a surrogate variable $Y = X + \xi$ where the pdf f_ξ of ξ is known. The real life applications of this model arise in econometrics, astronomy, biometrics, medical statistics, image reconstruction (see, e.g., Bovy et al. (2011) and also Carroll et al. (2006) and Meister (2009) and references therein). Density deconvolution problem was extensively studied in the last thirty years (see, e.g., Carroll et al. (2009), Comte and Kappus (2015), Goldenshluger (1999), Comte and Lacour (2011), and Meister (2009) and references therein). However, Berkson (1950) argued that in many situations it is more appropriate to treat the true unobserved variable as being contaminated with an error itself and search for the distribution of $W = X + \eta$ where η is the so-called Berkson error with a known pdf f_η. Here, X, ξ and η are assumed to be independent. The objective is to estimate the pdf f_W of W on the basis of i.i.d. observations

$$Y_i = X_i + \xi_i, \quad i = 1, \cdots, n, \tag{1.1}$$

where X_i and ξ_i are i.i.d with, respectively, the pdfs f_X which is unknown and f_ξ which is known. Density f_ξ is called the error (or the blurring) density. However, in majority of practical situations, the Berkson errors are small.

The objective of the first part of this dissertation is to study the situation where both the blurring and the Berkson errors are present and, in addition, the Berkson errors η_i, $i = 1, \cdots, n$, are small. To quantify this phenomenon, we assume that the pdf f_η is of the form $f_\eta(x) = \sigma^{-1} g\left(\sigma^{-1} x\right)$, where σ is small, specifically, $\sigma = \sigma_n \to 0$ as $n \to \infty$ while the variable X has a non-asymptotic scale. Specifically, we shall provide a full theoretical study of the bandwidth selection in a density deconvolution with Berkson errors including estimation construction and errors evaluation.

The second part of the dissertation is devoted to the study of network data. A network (graph) $G = (V, E)$ is a structure made of nodes V and edges E. The degree of a node represents the number of connections it has with other nodes in the network. Networks are more commonly represented as graphs as well as in terms of matrices known as adjacency matrices. An adjacency matrix is a binary matrix where $A_{ij} = 1$ if there is an edge connecting node i and node j, and

$A_{ij} = 0$ otherwise. In this dissertation, we consider an undirected network with n nodes and no self loops or multiple edges. We assume that elements of the adjacency matrix are generated as independent Bernoulli variables with P_{ij} being a probability of connections between nodes i and j so that $A_{ij} = \text{Ber}(P_{ij})$, where P is a symmetric probability matrix. That is, we perform one Bernoulli trial for each pair of edges, and record the result in both A_{ij} and A_{ji}.

A well-known feature of many empirical networks is community structure. Nodes in a network are often found to belong to groups or communities that exhibit similar behavior. In general, community structures may also refer to groups of vertices that connect in a similar manner to the rest of the nodes in a graph without having necessarily a higher inner density. For example, dis-assortative communities that have higher external connectivity. The primary interest in such networks is to understand which nodes exhibit similar behavior and in what way. The latter serves as the preliminary step towards other learning tasks. Some of the applications of community detection in networks include understanding of sociological behavior, protein to protein interaction, gene expressions, recommendation systems, medical prognosis, image segmentation, natural language processing, product-customer segmentation and web page sorting, to name a few.

The simplest model for the network is the Erdős Rényi random graph model $G(n, p)$ where n is the number of nodes and p is the probability of having an edge between two nodes, and all edges form independently with probability p. This model does not allow the community structure and is also too simplistic for the applications. The simplest random graph model for networks with community structure is the Stochastic Block Model (SBM). In this model the connection probability between two nodes is completely determined by the communities, to which the pair of nodes belong. For this reason, every nodes inside a community have same degree distribution and same expected degree, which is not true for majority of applications. The Degree Corrected Block Model (DCBM) introduces the node-dependent weights, so that the probability of connection between nodes inside the same community can vary. This makes DCBM more flexible than SBM. However, it still cannot realistically model the real life networks, since a node with a higher weight has uniformly higher connection probability in the network. The popularity of a node in a community k is defined as the expected number of connection of that node to the nodes in the community k. The Popularity Adjusted Block Model (PABM) recently introduced in Sengupta and Chen (2018), defines the probability of connection between nodes as the product of the nodes' popularities in the communities where another node belongs.

In this dissertation, we focus on community detection and estimation of the probability of connections matrix in the PABM. We develop methodology for estimation of P and clustering of nodes into communities under this model. In addition, we provide non-asymptotic guarantees for the estimation and clustering errors. One of the advantages of our approach is that we do not assume that the number of communities in the network is known. Our estimation procedure provides the oracle upper bounds for the error by imposing a penalty on the unknown number of communities. Since the majority of real-life networks are sparse, we introduce the sparse PABM model. The majority of sparse network models in the literature are based on a rather unrealistic assumption that the maximum connection probability is bounded above by a small quantity. The reason for this is that the SBM and DCBM do not allow to set probability of connection between two nodes equal to zero since it is equal to the probability of connection between the communities to which they belong, or make the nodes disconnected from the network. The flexibility of the PABM allows modeling some of the probabilities of connections between the nodes as identical zeros while maintaining the rest of the probabilities non-negligible, leading to the Sparse Popularity Adjusted Block Model (SPABM). This formulation reduces the size of parameter set and leads to improved precision of estimation and clustering. We produce the estimators of the probability matrix and the community structure in this setting. Furthermore, we provide non-asymptotic upper bounds for the estimation and the clustering errors.

The rest of the dissertation is organized as follows.

In Chapter 2, we provide the complete theoretical treatment of density deconvolution with small Berkson errors. Section 2.1 consists of background information that are used frequently in the dissertation. We start with the brief development of density estimation techniques from histogram to kernel estimator in Section 2.1.1. Section 2.1.2 contains the definition and properties of Fourier transform and we present the important probabilistic tail inequalities that were used in proving our results in Section 2.1.3. We formulate the deconvolution problem in Section 2.2. Section 2.3 constructs the deconvolution estimator. We evaluate the estimation error in Section 2.4. In particular, it provides an oracle inequality for the risk and studies the upper bounds for the risk under specific assumptions on the class of underlying functions. We talk about adaptive estimation with the Lepski's method in Section 2.5. Section 2.6 provides the discussion over the results and the proofs of the main and supplementary results can be found in Section 2.7.

Chapter 3 considers the estimation and clustering in the Popularity Adjusted Stochastic Block Model. Section 3.1 briefly describes the background of statistical network models. Section 3.1.1 provides a brief description of random networks. We introduce the community structure and describe the Stochastic Block Model, Degree Corrected Block Model and Popularity Adjusted Stochastic Block Model in the Section 3.2. Section 3.3 describes the structure of the probability matrix and the advantage of PABM in modeling network data. Section 3.4 introduce the notations we used for the second part of the dissertation. The optimization procedure to estimate the probability matrix and clustering matrix is discussed in Section 3.5. We compute the estimation and clustering errors in Section 4.3. Finally, the proofs of all the main results and supplementary results for estimation and clustering in PABM are available in Section 3.7.

Chapter 4 is devoted to the estimation and clustering in the Sparse Popularity Adjusted Blockmodel. Section 4.1 introduces definitions of sparsity in blockmodels and discuss on the advantage of PABM in modeling sparsity. We propose the optimization procedure for estimation and clustering in sparse PABM in Section 4.2, where Section 4.2.1 discuss the structure of the probability matrix in sparse PABM and Section 4.2.2 formulate the optimization problem for estimation and clustering in sparse PABM. We compute the estimation and clustering errors in sparse PABM in Section 4.3. We talk about two different types of penalty and its expression in the Section 4.3.1. Our main results on the estimation and clustering errors are presented in the Section 4.3.2 and Section 4.3.3 respectively. The proof of the main and supplementary results in sparse PABM can be found in Section 4.4.

In Chapter 5, we discuss on the results obtained in the PABM and the sparse PABM. In addition, we briefly talk about the possible direction of the future work.

CHAPTER 2: DENSITY DECONVOLUTION WITH BERKSON ERRORS

The results presented in this chapter starting from the section 2.2 have been published in Rimal and Pensky (2019).

2.1 Background Material

In this section we provide the background material that is required to comfortably read the dissertation. In this regards we define the terminology that were used in the dissertation. In addition, we state the results that will be used later.

2.1.1 Introduction to Density Estimation

Let $X_1, \cdots, X_n$ be identically distributed real valued random variables. Let the common distribution of i.i.d random vaariables $X_1, \cdots, X_n$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$. Then the density of this distribution is a unknown function f such that $f : \mathbb{R} \to [0, \infty)$ which we wanted to estimate. The density estimation is a construction of estimate of density function from the observed data.

For a random variable X, with probability density function (pdf) f, one has

$$\mathbb{P}(a < X < b) = \int_a^b f(x)dx \quad \text{for all} \quad a < b.$$

An estimator of f is a function $x \mapsto f_n(x, X_1, \cdots, X_n)$ measurable with respect to the observation $(X_1, \cdots, X_n)$. There are two approaches to density estimation, which are Parametric Density Estimation and Non Parametric Density Estimation. Suppose we know that f belongs to a parametric family $g(x, \theta) : \theta \in \Theta$, where g is a given function, and $\Theta \subset \mathbb{R}^p$ with a fixed dimension p independent of n. Then the problem of estimation of f is equivalent to estimation of the parameter θ which is known as parametric problem of density estimation. In the parametric density estimation, we know the shape of the distribution but we don't know the values of the parameter. We usually estimate the values of the parameter by using Maximum Likelihood Estimation methods or Bayesian Estimation methods.

For example: suppose the n data points $X_1, X_2, \cdots, \quad X_n$ are observed. Assume that

$$X_1, X_2, \cdots, \quad X_n \sim \mathcal{N}(\mu, \sigma^2)$$

Then we estimate μ and σ^2 from the data, we usually estimate the values of the parameter by using Maximum Likelihood Estimation methods or Bayesian Estimation methods. Finally

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

will be the pdf for the given data.
The main limitation of the parametric density estimation is that density function of the data were constrained to fall in a given parametric family.

On the other hand, if we don't have information about f then our problem of estimating f becomes the non-parametric density estimation problem. In this case, we assume that f belongs to some wide class of densities (also known as nonparametric classes of functions) $\mathcal{F}$. For example, $\mathcal{F}$ can be the Holder class of densities, or Sobolev class of densities, or set of all the continuous probability densities on $\mathbb{R}$ or the set of all the Lipschitz continuous probability densities on $\mathbb{R}$ and many more. In the nonparametric density estimation, the form of the density is entirely determined by the data without any model. Our work is mainly related to the non parametric density estimation for the univariate case. We start with the review of some common methods of density estimation.

Histogram

Histogram is an oldest and most widely used density estimator. Given the data $X_0, X_1, X_2, \cdots, X_n$, First we specify the origin c_0, then define

$$a = \min_{0 \le i \le n} X_i, \quad b = \max_{0 \le i \le n} X_i$$

and divide $[a, b]$ into k intervals of equal length $h = \dfrac{b-a}{k}$, then

$$I_j = (c_0 + jh, c_0 + (j+1)h], \quad j = \cdots, -1, 0, 1, \cdots$$

We choose the intervals open on the left and closed on the right for definiteness. Then the estimator of the density is given by

$$\hat{f}(x) = \frac{1}{nh} \sum_i \mathbb{I}(X_i \in I_j), \quad x \in I_j$$

The form of the histogram depends on the tuning parameters c_0 and h.

The major drawback of the histogram is the discontinuity. We cannot use for further mathematical treatment. A histogram is also difficult to construct as well as visualize in the multivariate case because it is not easy to construct contour diagrams to represent the data. The problems in the univariate case are worsened in the multivariate case because of the dependence of estimates on the choice of origin and coordinate directions of the grid of cells.

The Naive Estimator

For any given h, we can estimate $\mathbb{P}[x \in (x-h, x+h)]$ by the proportion of the sample falling in the interval $(x-h, x+h)$.

The natural estimator $\hat{f}$ of the density is given by choosing small number h with $\hat{f}(x)$ equals the proportion of X_i's falling in the interval $(x-h, x+h)$ divided by the n. That is

$$\hat{f}(x) = \frac{1}{2hn} \sum_i \mathbb{I}[X_i \in (x-h, x+h)]$$

which can be written more formally as

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=0}^{n} w\left(\frac{x - X_i}{h}\right)$$

where weight function is defined by

$$w(x) = \begin{cases} 1/2 & if \quad |x| < 1 \\ 0 & otherwise. \end{cases}$$

Naive estimator $\hat{f}(x)$ can be seen as constructing a histogram with every point is the center of the sampling interval and $\hat{f}(x)$ is the ordinate of histogram at x. The density function obtained is not continuous but piecewise constant with zero derivatives everywhere else. This makes it less attractive as it cannot be used for further mathematical treatment where we need to deal with the differentiation.

The Kernel Estimator

The kernel function $K : \mathbb{R} \to \mathbb{R}$ is an integrable function satisfying $\int_{-\infty}^{\infty} K(z)dz = 1$. The Kernel Estimator is a generalization of the naive estimator by replacing the weight function w with a kernel function K. The Kernel density estimator $\hat{f}$ of the density f is given by

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=0}^{n} K\left(\frac{x - X_i}{h}\right)$$

where K is a kernel function and the parameter h is called a bandwidth of the estimator. The Kernel estimator can be considered as the sum of the bumps placed at the observations whereas the naive estimator is a sum of boxes centered at the observations. The Kernel estimator is the most commonly studied and widely used density estimator. If the kernel function is non-negative everywhere, then the estimator will be a probability density which inherits the continuity and differentiability property of the kernel K. Although Kernel estimator suffers from the tendency of appearing spurious noise in the tails of the estimates when applied to data from long-tailed distributions, there have been various adaptive methods available in the literature (for example: look Silverman (1986)) to deal with the issue.

2.1.2 Fourier Transform

Definition 2.1.1. *For any function $f \in L_1(\mathbb{R})$, we define its Fourier transform f^* by*

$$f^*(x) = \int_{-\infty}^{\infty} e^{ixt} f(t)dt$$

Definition 2.1.2. *Let X and Y be two continuous random variables with density functions f and g respectively defined on $\mathbb{R}$. Then the convolution of f and g is denoted by $f * g$ is a real valued function defined by*

$$(f * g)(z) = \int f(z - y)g(y)dy = \int f(x)g(z - y)dx = (g * f)(z)$$

The distribution of the random variable X is the information about what values X takes with what probabilities. The distribution of X is determined by the cumulative distribution function (CDF) defined by

$$F_X(t) = \mathbb{P}(X \leq t),\ t \in \mathbb{R}$$

The relation

$$\mathbb{P}(X > t) = 1 - F_X(t)$$

is useful to work with the tails of random variables.

The following lemma shows the density of sum of the two independent variables is equal to the convolution of their densities.

Lemma 2.1.1. *Let X and Y be two independent, continuous random variables with density functions f_X and f_Y respectively defined on $\mathbb{R}$. Then the sum $Z = X + Y$ is a continuous random variable with density function f_Z given by $f_Z = f_X * f_Y$.*

Proof. Let X and Y be two independent, continuous random variables with density functions f_X and f_Y respectively defined on $\mathbb{R}$. Then the sum $Z = X + Y$ is a continuous random variable with the cumulative density function (cdf)

$$\begin{aligned} F_{X+Y}(z) &= \mathbb{P}(X + Y \leq z) \\ &= \int_{x+y\leq z} f_X(x) f_Y(y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f_X(x) f_Y(y) dy dx \\ &= \int_{-\infty}^{\infty} F_Y(z-x) f_X(x) dx \end{aligned}$$

Then the density function of $Z = X + Y$ is given by

$$\begin{aligned} f_Z(z) &= \frac{d}{dz}\big(F_{X+Y}(z)\big) \\ &= \int_{-\infty}^{\infty} \frac{d}{dz} F_Y(z-x) f_X(x) dx \\ &= \int_{-\infty}^{\infty} f_Y(z-x) f_X(x) dx \\ &= (f_X * f_Y)(z) \end{aligned}$$

□

The following lemma shows that Fourier transform takes convolution to the product.

Lemma 2.1.2. *(Convolution theorem) Let f^* and g^* are the Fourier transforms of a function f and and g defined by*

$$f^*(x) = \int_{-\infty}^{\infty} e^{ixt} f(t) dt, \quad \textit{and} \quad g^*(x) = \int_{-\infty}^{\infty} e^{ixt} g(t) dt$$

For any $f, g \in L_1(\mathbb{R})$, the Fourier transform of the convolution $f * g$ is given by

$$(f * g)^*(x) = f^*(x) g^*(x)$$

Plancherel theorem shows the relation between the integral of a function f to the integral of its Fourier transform.

Lemma 2.1.3. *(Plancherel theorem) For any $f \in L_1(\mathbb{R})$, let f^* denote the Fourier transforms of a function f defined by*

$$f^*(x) = \int_{-\infty}^{\infty} e^{ixt} f(t) dt$$

Then for any $f \in L_1(\mathbb{R}) \cap L_2(\mathbb{R})$

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |f^*(w)|^2 dw$$

The following lemma provides the condition for the existence of the inverse Fourier transform.

Lemma 2.1.4. *Assume that $f \in L_1(\mathbb{R})$ is bounded and continuous at some $x \in \mathbb{R}$ and, in addition, $f^* \in L_1(\mathbb{R})$. Then*

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ixt} f^*(t) dt$$

For the proof of the above lemma, refer to page 181 of Meister (2009).

2.1.3 Probabilistic Tail Inequalities

Concentration inequalities quantify how a random variable X deviates around its mean μ. They form a two-sided bounds for the tails of $X - \mu$ such as

$$\mathbb{P}(|X - \mu| > t) \leq \exp(-f(t))$$

where $f(t)$ is some increasing function of t.

Lemma 2.1.5. *(Markov Inequality) Let X be a non negative random variable and suppose that $\mathbb{E}(X)$ exists. Then for any positive real number t,*

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(X)}{t}$$

The proof of the lemma can be found in page 8 of the book High-Dimensional Probability by

Vershynin (2018).

Lemma 2.1.6. *(Hoeffding's inequality) Let $X_1, \cdots, X_n$ be independent random variables such that $a_i \leq X_i \leq b_i$. Then for all $t > 0$*

$$\mathbb{P}\left(\sum_{i=1}^{n}(X_i - \mathbb{E}(X_i)) \geq t\right) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right)$$

The proof of the lemma can be found in page 14 of the book High-Dimensional Probability by Vershynin (2018).

Lemma 2.1.7. *(Theorem 2.1, Hsu et al. (2012)) Let $A \in \mathbb{R}^{n\times n}$ be a matrix, and let $\Sigma = Tr(A^T A)$. Suppose that $x = (x_1, \cdots, x_n)$ is a random vector such that, for some $\mu \in \mathbb{R}^n$ and $\sigma \geq 0$,*

$$\mathbb{E}[\exp(\alpha^T(x - \mu))] \leq \exp\left(\|\alpha\|^2\sigma^2/2\right)$$

for all $\alpha \in \mathbb{R}^n$. For all $t > 0$

$$\mathbb{P}\left\{\|Ax\|^2 \geq \sigma^2(Tr(\Sigma) + 2\sqrt{Tr(\Sigma^2)\,t} + 2\|\Sigma\|t) + Tr(\Sigma\mu\mu^T)\left(1 + 2\left(\frac{\|\Sigma\|^2}{Tr(\Sigma^2)}t\right)^{1/2}\right)\right\} \leq \exp(-t).$$

The proof of the above result can be found in page 3 of the article Hsu et al. (2012).

Definition 2.1.3. *(Sub-gaussian random variables) A random variable X that satisfies one of the three equivalent properties given below is called a sub-gaussian random variable.*

1. Tails: $\mathbb{P}\{|X| > t\} \leq \exp\left(1 - t^2/C_1^2\right)$ for all $t \geq 0$;

2. Moments: $(\mathbb{E}|X|^p)^{1/p} \leq C_2\sqrt{p}$ for all $p \geq 1$;

3. Super-exponential moment: $\mathbb{E}\exp(X^2/C_3^2) \leq e$, where the parameters C_1, C_2, C_3 differing *from each other by at most an absolute constant factor.*

Nonasymptotic bound on the norm of a random matrix is very useful on various areas of pure and applied mathematics (see Davidson and Szarek (2001), Rudelson and Vershynin (2010), Talagrand (2014)). Bandeira and van Handel (2016) are interested to obtain upper and lower bounds on $\|X\|$ in terms of natural parameters that capture the structure of X, that differ only by universal constants. They consider investigating a norm of random matrices with independent entries. Let X be an $n \times n$ symmetric matrix with $X_{ij} = g_{ij}b_{ij}$, where $\{g_{ij} : i \geq j\}$ are $i.i.d \sim N(0, 1)$ and $\{b_{ij} : i \geq j\}$ are given scalers. In this setting, the most useful non-asymptotic bound on the spectral norm $\|X\|$ due to consequence of non-commutative Khintchine inequality (see Lust-Piquard and Pisier (1991)) yields

$$\mathbb{E}\|X\| \lesssim \sigma\sqrt{\log(n)}, \quad \sigma = \max_i \sqrt{\sum_j b_{ij}^2},$$

which fails to be sharp in the case of Wigner matrices. Since $\sigma = \sqrt{n}$ results to

$$\mathbb{E}\|X\| \lesssim \sqrt{n}\sqrt{\log(n)}$$

lacks correct scaling $\mathbb{E}\|X\| \sim \sqrt{\log(n)}$. But if we take X a diagonal matrix with independent standard gaussian entries, then $\sigma = 1$ and hence $\mathbb{E}\|X\| \sim \sqrt{\log(n)}$. So the non-commutative Khintchine bound is sharp in extreme cases, it fails to capture the structure of the matrix X in a satisfactory manner.

Bandeira and van Handel (2016) proved the following results.

Lemma 2.1.8. *(Theorem 1.1: Bandeira and Handel (2016)) Let X be an $n \times n$ symmetric matrix with $X_{ij} = g_{ij}b_{ij}$, where $\{g_{ij} : i \geq j\}$ are $i.i.d \sim N(0, 1)$ and $\{b_{ij} : i \geq j\}$ are given scalers. Then*

$$\mathbb{E}\|X\| \leq (1 + \varepsilon)\left\{2\sigma + \frac{6}{\log(1 + \varepsilon)}\sigma_* \sqrt{\log(n)}\right\}$$

for any $0 \le \varepsilon \le 1/2$*, where* σ*, and* σ_* *are defined by*

$$\sigma = \max_i \sqrt{\sum_j b_{ij}^2}, \quad \textit{and } \sigma_* = \max_{i,j} |b_{ij}|.$$

Let X be an $n \times m$ matrix with $X_{ij} = g_{ij} b_{ij}$, where $\{g_{ij} : 1 \le i \le n, 1 \le j \le m\}$ are independent $N(0, 1)$ random variables and $\{b_{ij} : 1 \le i \le n, 1 \le j \le m\}$ are given scalers. One can obtain a bound on $\mathbb{E}\|X\|$ for a non symmetric matrix X by applying Theorem 1 on the symmetric matrix

$$\tilde{X} = \begin{bmatrix} 0 & X \\ X^* & 0 \end{bmatrix}$$

They obtain the following result for rectangular matices.

Lemma 2.1.9. *(Theorem 3.1: Bandeira and Handel (2016)) Let* X *be an* $n \times m$ *matrix with* $X_{ij} = g_{ij} b_{ij}$*. Then*

$$\mathbb{E}\|X\| \le (1 + \varepsilon) \left\{ \sigma_1 + \sigma_2 + \frac{5}{\log(1 + \varepsilon)} \sigma_* \sqrt{\log(n \wedge m)} \right\}$$

for any $0 \le \varepsilon \le 1/2$*.*

Then they extend the result for independent random variables which is as follows.

Lemma 2.1.10. *(Corollary 3.3: Bandeira and Handel (2016)) If the independent Gaussian variables* g_{ij} *are replaced by independent random variables* ξ_{ij} *that are centered and sub-Gaussian in the sense*

$$\mathbb{E}[\xi_{ij}] = 0, \quad \mathbb{P}[|\xi_{ij} > t|] \le C \exp\{-t^2/2c\} \quad \textit{for all } t \ge 0, \textit{ and } i, j;$$

then the Theorems 1.1 and 3.1 remain valid up to a universal constant that depends on C and c only. That is, we have $\mathbb{E}\|X\| \lesssim \sigma + \sigma_* \sqrt{\log(n)}$ in the case of Theorem 1.1 and $\mathbb{E}\|X\| \lesssim \sigma_1 + \sigma_2 + \sigma_* \sqrt{\log(n \wedge m)}$ in the case of Theorem 3.1.

Lemma 2.1.11. *(Singular value decomposition) Let A be a matrix of size $n \times p$. Let r be the rank of the matrix A. Then A can be decomposed as*

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

where

(i) $\sigma_1 \geq \sigma_2, \cdots, \geq \sigma_r > 0$, *and* $\sigma_1, \sigma_2, \cdots, \sigma_r$ *are called the singular values of* A

(ii) $\sigma_1^2 \geq \sigma_2^2, \cdots, \geq \sigma_r^2$ *are the nonzero eigenvalues of* $A^T A$ *and* AA^T, *and*

(iii) $\{u_1, \cdots, u_r\}$ *and* $\{v_1, \cdots, v_r\}$ *are two orthonormal families of* $\mathbb{R}^n$ *and* $\mathbb{R}^p$ *such that* $AA^T u_i = \sigma_i^2 u_i$ *and* $A^T A v_i = \sigma_i^2 v_i$ *for* $i = 1, 2, \cdots, r$.

The proof of the SVD can be found in page 229 of the Giraud (2015).

Definition 2.1.4. *(Frobenius norm)*

The standard scalar product on matrices is $\langle A, B \rangle_F = \sum_{i,j} A_{ij} B_{ij}$. It induces the Frobenius norm

$$\|A\|_F^2 = Tr(A^T A) = \sum_{i=1}^{r} \sigma_i^2$$

where r is the rank of a matrix A.

Definition 2.1.5. *(Operator norm) The operator norm of a matrix A is defined by $\|A\|_{op}^2 = \sigma_1$ where σ_1 is the largest singular value of a matrix A.*

Definition 2.1.6. *(Nuclear norm) The nuclear norm of a matrix A is defined by*

$$\|A\|_* = \sum_{i=1}^{r} \sigma_i$$

where $\sigma_1, \sigma_2, \cdots, \sigma_r$ are the singular values of A.

Definition 2.1.7. *(Ky-Fan (2,q) norm) For any integer $q \geq 1$, the Ky-Fan (2,q) norm of a matrix A is defined by*

$$\|A\|^2_{(2,q)} = \sum_{i=1}^{q} \sigma_i^2$$

where $\sigma_1, \sigma_2, \cdots, \sigma_r$ are the singular values of A.

We have the following relation between the matrix norms defined above.

$$\|A\|_* \leq \sqrt{\text{rank}(A)}\|A\|_F;\ \langle A, B\rangle_F \leq \|A\|_*\|A\|_F, and\ \|AB\|_F \leq \|A\|_{op}\|B\|_F$$

$$|A\|_{(2,1)} = \|A\|_{op} = \sigma_1$$

Definition 2.1.8. *(Moore-Penrose pseudo-inverse of arbitrary matrix A) Let $A = \sum_{i=1}^{r} \sigma_i u_i v_i^T$ be a singular value decomposition of A with $r =$ rank(A). The Moore-Penrose pseudo-inverse of A is given by*

$$A^+ = \sum_{i=1}^{r} \sigma_i{}^{-1} v_i u_i^T$$

The problem of predicting m-dimensional vector y from a p-dimensional vector of covariates can be formulated in a matrix form as

$$Y = XA_* + E$$

where $Y = [y_k^{(i)}], i = 1, 2, \cdots, n; k = 1, 2, \cdots, m$ and $E = [\varepsilon_k^{(i)}], i = 1, 2, \cdots, n; k = 1, 2, \cdots, m.$ and $X = [x_j^{(i)}], i = 1, 2, \cdots, n; j = 1, 2, \cdots, p$. If we know the rank r_* of A_*.

Then we would estimate A_* by the maximum likelihood estimator $\hat{A}_r$ constrained to have a rank at most r_* by

$$\hat{A}_r \in \underset{\text{rank}(A) \leq r}{argmin} \|Y - XA\|_F^2$$

with $r = r_*$. The following lemma provides a useful formula for $X\hat{A}_r$ in terms of the singular value decomposition of PY, where P is the orthogonal projector onto the range of X.

Lemma 2.1.12. *Write $P = X(X^TX) + X^T$ for the projection onto the range of X, with $(X^TX)_+$ the Moore-Penrose pseudo-inverse of X^TX. Then, for any $r \geq 1$, we have $X\hat{A}_r = (PY)_r$. As a consequence, denoting by $PY = \sum_{k=1}^{rank(PY)} \sigma_k u_k {v_k}^T$ a singular value decomposition of PY, we have for any $r \geq 1$*

$$X\hat{A}_r = \sum_{k=1}^{r \wedge rank(PY)} \sigma_k u_k {v_k}^T$$

Lemma 2.1.13. *(Proposition 6.2 Giraud (2015))*

Set rank$(X) = q$. For any $r \geq 1$ and $\theta > 0$, we have

$$\|X\hat{A}_r - XA^*\|_F^2 \leq C^2(\theta) \sum_{k>r} \sigma_k(XA^*)^2 + 2C(\theta)(1+\theta)r|PE|_{op}^2$$

with $C(\theta) = 1 + \frac{2}{\theta}$.

The proof of the above lemmas can be found in page 124 of the Giraud (2015).

2.2 Deconvolution Problem with Berkson Errors: Formulation

Estimation with Berkson errors occurs in a variety of statistics fields such as analysis of chemical, nutritional, economics or astronomical data (see, e.g., Kim et al. (2016), Long et al. (2016), Robinson (1999), Wang (2003), Wason et al. (1984), among others). For example, in occupational medicine, an important problem is the assessment of the health hazard of specific harmful substances in a working area. A modeling approach usually assumes that there is a threshold concentration, called the threshold limiting value (TLV), under which there is no risk due to the substance. Estimating the TLV is of particular interest in the industrial workplace. The classical errors in this model come from the measures of dust concentration in factories, while the Berkson errors come from the usual occupational epidemiology construct, wherein no direct measures of dust exposure are taken on individuals, but instead plant records of where they worked and for how long are used to impute some version of dust exposure (see Carroll et al. (2006)). In economics, the household income is usually not precisely collected due to the survey design or data sensitivity. It was described by Kim et al. (2016) (see also Geng and Koul (2018)) that when the income data were collected by asking individuals which salary range categories they belong to, then the midpoint of the range interval was used in analysis. In this case, it is wise to assume that the true income fluctuates around the midpoint observation subject to errors.

Estimation with Berkson errors was studied by Carroll et al. (2009), Delaigle (2007), Delaigle (2008), Du et al. (2011), Geng and Koul (2018), Wang (2003), Wang et al. (2004) among others. It is well known that the presence of Berkson errors improves precision of estimation of the density function f_W. For example, Delaigle (2007), Delaigle (2008) who studied estimation with Berkson errors noted that in the cases when the pdf f_η of Berkson errors has higher degree of smoothness than the error density f_ξ, one can obtain estimators of f_W with the parametric convergence rate.

However, in majority of practical situations, the Berkson errors are small. Hence, the question arises whether small Berkson errors improve the estimation accuracy and how much. A similar inquiry has been recently carried out by Long et al. (2016) who considered a somewhat different setting. In particular, they studied a p-dimensional version of the problem where variable X is directly observed and the objective is estimation of the pdf f_W of $W = X + \eta$ on the basis of observations $X_1, \cdots, X_n$ where the pdf f_η of η is known and variable η is small. In this formulation, the pdf f_W can be written as

$$f_W(x) = \int_{\mathbb{R}^p} f_X(x-z) f_\eta(z) dz$$

and can be estimated by

$$\hat{f}_W(x) = n^{-1} \sum_{i=1}^{n} f_\eta(x - X_i) \tag{2.1}$$

with the parametric error rate of Cn^{-1}. However, if Var$(\eta) = \sigma^2$ is small, this rate becomes $C(\sigma)n^{-1}$ where $C(\sigma) \to \infty$ when $\sigma \to 0$, so the error of the estimator (2.1) may be very high.

To resolve this difficulty, in addition to estimator (2.1), Long et al. (2016) proposed two alternative kernel estimators where the bandwidths of the kernels are chosen as $h = h_W$ or $h = h_X$, so to minimize the error of the estimator of f_W in the first case and the error of the estimator of f_X in the second case. Subsequently, the authors studied all three estimators by simulations and concluded that overall the kernel estimator with $h = h_W$ outperforms the remaining two. When the error variance σ is small, the estimator (2.1) leads to sub-optimal error rates. On the other hand, the choice of $h = h_X$ leads to oversmoothing, especially, when the error variance is large. The authors do not provide a comprehensive theoretical study of the bandwidth selection in a general case. In particular, their rule-of-thumb recipe is based on the case where f_X is a Gaussian density. In particular, Long et al. (2016) did not investigate when estimator (2.1) that corresponds to the bandwidth $h = 0$ is preferable and suggested that it is always suboptimal.

The objective of the dissertation is to study the situation where both the blurring and the Berkson errors are present and, in addition, the Berkson errors η_i, $i = 1, \cdots, n$, are small. To quantify this phenomenon, we assume that the pdf f_η is of the form

$$f_\eta(x) = \sigma^{-1} g\left(\sigma^{-1} x\right), \tag{2.2}$$

where σ is small, specifically, $\sigma = \sigma_n \to 0$ as $n \to \infty$ while the variable X has a non-asymptotic scale. Specifically, we shall provide a full theoretical study of the bandwidth selection in a density deconvolution with Berkson errors.

The setting of Long *et al.* (2016) corresponds to the multivariate version of the problem in this paper where $\xi_i = 0$ and $f_\xi^* = 1$. We provide full theoretical treatment of the problem. In particular, we prove that one should always choose the bandwidth to minimize the error of the estimator of f_W, but in some cases this optimal bandwidth can be zero if σ lies above some threshold that depends on the shapes of the densities f_ξ, f_X and g and the sample size. In the particular case studied by Long et al. (2016) the latter situation would lead to the estimator of the form (2.1).

Since the setting (2.2) leads to three asymptotic parameters, n, σ and h, in order to keep the paper clear and readable, we consider a one-dimensional version of the problem. Extensions of our results to the situation of multivariate densities is a matter of future work.

In what follows, we are using the following notations. For any function f, f^* denotes its Fourier transform defined by $f^*(x) = \int_{-\infty}^{\infty} e^{ixt} f(t)dt$. If f is a pdf, then f^* is the characteristic function of f. We use the symbol C for a generic positive constant, which takes different values at different places and is independent of n. Also, for any positive functions $a(n)$ and $b(n)$, we write $a(n) \asymp b(n)$ if the ratio $a(n)/b(n)$ is bounded above and below by finite positive constants independent of n, and $a(n) \lesssim b(n)$ if the ratio $a(n)/b(n)$ is bounded above by finite positive constants independent of n.

2.3 Construction of the Deconvolution Estimator

Since (1.1) and $W = X + \eta$ imply that

$$f_Y^*(w) = f_X^*(w) f_\xi^*(w), \quad f_W^*(w) = f_X^*(w) f_\eta^*(w) \tag{2.3}$$

and also, due to (2.2), $f_\eta^*(w) = g^*(\sigma w)$, one obtains

$$f_W^*(w) = f_X^*(w) g^*(\sigma w) = \frac{f_Y^*(w) g^*(\sigma w)}{f_\xi^*(w)}$$

Note that the unbiased estimator of $f_Y^*(w)$ is given by the empirical characteristic function

$$\hat{f}_Y^*(w) = n^{-1} \sum_{j=1}^{n} \exp(iwY_j). \tag{2.4}$$

If $g^*(\sigma w)/f_\xi^*(w)$ is square integrable, i.e.

$$\rho^2(\sigma) = \int_{-\infty}^{\infty} \left| \frac{g^*(\sigma w)}{f_\xi^*(w)} \right|^2 dw < \infty, \tag{2.5}$$

then the inverse Fourier transform of $f_Y^*(w) g^*(\sigma w)/f_\xi^*(w)$ exists and $f_W(x)$ can be estimated by

$$\hat{f}_W(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-iwx) \, \frac{\hat{f}_Y^*(w) g^*(\sigma w)}{f_\xi^*(w)} dw \tag{2.6}$$

If $g^*(\sigma w)/f_\xi^*(w)$ is not square integrable, one needs to obtain a kernel estimator of f_W. Construct approximations $f_{W,h}$ and $f_{W,h}^*$ of f_W and f_W^*, respectively,

$$f_{W,h}(x) = \int_{-\infty}^{\infty} \frac{1}{h} K\left(\frac{x-w}{h}\right) f_W(w) dw, \quad f_{W,h}^*(s) = K^*(sh) \frac{f_Y^*(s) g^*(\sigma s)}{f_\xi^*(s)} \tag{2.7}$$

and arrive at the estimator $\hat{f}^*_{W,h}(s)$ of $f^*_{W,h}(s)$ of the form

$$\hat{f}^*_{W,h}(s) = K^*(sh)\hat{f}^*_Y(s)g^*(\sigma s)/f^*_\xi(s)$$

where $\hat{f}^*_Y$ is defined in (2.4).

Consider the kernel function $K(x) = \sin(x)/(\pi x)$, so that $K^*(s) = I(|s| \leq 1)$, where $I(A)$ denotes the indicator function of a set A. Since $K^*(s)$ is bounded and compactly supported, the inverse Fourier transform of $\hat{f}^*_{W,h}$ always exists and

$$\hat{f}_{W,h}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-ixs)\,\frac{\hat{f}^*_Y(s)K^*(sh)g^*(\sigma s)}{f^*_\xi(s)}\,ds \tag{2.8}$$

We set $\hat{f}_{W,0}(x) \equiv \hat{f}_W(x)$.

In order to obtain an expression for the bandwidth h we introduce the following assumptions:

(A1) There exists positive numbers c_ξ and C_ξ and nonnegative numbers a, b and d such that for any s

$$c_\xi(s^2+1)^{-\frac{a}{2}}\exp(-d|s|^b) \leq |f^*_\xi(s)| \leq C_\xi(s^2+1)^{-\frac{a}{2}}\exp(-d|s|^b) \tag{2.9}$$

where $b = 0$ iff $d = 0$ and $a > 0$ whenever $d = 0$.

(A2) There exists positive numbers c_g and C_g and nonnegative numbers ϑ, β and γ such that for any s

$$c_g(s^2+1)^{-\frac{\vartheta}{2}}\exp(-\gamma|s|^\beta) \leq |g^*(s)| \leq C_g(s^2+1)^{-\frac{\vartheta}{2}}\exp(-\gamma|s|^\beta), \tag{2.10}$$

where $\beta = 0$ iff $\gamma = 0$ and $\vartheta > 0$ whenever $\gamma = 0$.

(A3) $f_X(s)$ belongs to the Sobolev ball

$$\mathcal{S}(k, B) = \{f : \int_{-\infty}^{\infty} |f_X^*(s)|^2 (s^2+1)^k \, ds \le B^2, k \ge 1/2\}. \tag{2.11}$$

Also, since density deconvolution with Berkson errors of relatively large size has been fairly well studied, below we only study the case where σ is small, in particular, if $\gamma > 0$, $d > 0$, one has

$$\sigma < 0.5 \, (d/\gamma)^{1/b}. \tag{2.12}$$

2.4 Estimation Error

We characterize the precision of the estimator $\hat{f}_{W,h}$ of f_W by its Mean Integrated Squared Error (MISE)

$$\text{MISE}(\hat{f}_{W,h}, f_W) = \mathbb{E} \int_{-\infty}^{\infty} |\hat{f}_{W,h}(x) - f_W(x)|^2 dx.$$

Since, under Assumptions (2.9)–(2.11), both $\hat{f}_{W,h}^*$ and f_W^* are square integrable, by the Plancherel theorem, derive that

$$\text{MISE}(\hat{f}_{W,h}, f_W) = \frac{1}{2\pi} \mathbb{E} \int_{-\infty}^{\infty} \frac{|g^*(\sigma s)|^2}{|f_\xi^*(s)|^2} \, |K^*(sh) \hat{f}_Y^*(s) - f_Y^*(s)|^2 \, ds$$

Therefore,

$$\text{MISE}(\hat{f}_{W,h}, f_W) = R_1(\hat{f}_{W,h}, f_W) + n^{-1} \, R_2(\hat{f}_{W,h}, f_W) \tag{2.13}$$

where

$$R_1(\hat{f}_{W,h}, f_W) = \|\mathbb{E}\hat{f}_{W,h} - f_W\|^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} |g^*(\sigma s)|^2 \, |f_X^*(s)|^2 \, I(|s| > h^{-1}) \, ds \tag{2.14}$$

is the integrated squared bias of the estimator $\hat{f}_{W,h}$ and

$$R_2(\hat{f}_{W,h}, f_W) = n \, \mathbb{E}\|\hat{f}_{W,h} - \mathbb{E}\hat{f}_{W,h}\|^2 \le I(\sigma, h) \tag{2.15}$$

where

$$I(\sigma, h) = \frac{1}{2\pi} \int_{-1/h}^{1/h} \frac{|g^*(\sigma s)|^2}{|f_\xi^*(s)|^2} \, ds. \tag{2.16}$$

We shall be interested in the maximum value of MISE$(\hat{f}_{W,h}, f_W)$ over all $f_X \in \mathcal{S}(k, B)$ where $\mathcal{S}(k, B)$ is defined in (2.11). In particular, we denote $\mathbb{E}\hat{f}_{W,h} = f_{W,h}$ and define

$$\Delta \equiv \Delta(n, \sigma, h) = \max_{f_X \in \mathcal{S}(k,B)} \text{MISE}(\hat{f}_{W,h}, f_W) \quad \text{subject to} \quad f_W^*(w) = f_X^*(w) f_\eta^*(w). \tag{2.17}$$

It is easy to see that

$$\Delta \le \Delta_1 + n^{-1} \, \Delta_2 \tag{2.18}$$

where

$$\Delta_1 \equiv \Delta_1(n, \sigma, h) = \max_{f_X \in \mathcal{S}(k,B)} R_1(\hat{f}_{W,h}, f_W), \quad \Delta_2 \equiv \Delta_2(n, \sigma, h) = \max_{f_X \in \mathcal{S}(k,B)} R_2(\hat{f}_{W,h}, f_W) \tag{2.19}$$

Then, the following statements hold.

Lemma 2.4.1. *Under the assumptions* (2.9)–(2.12)*, for* Δ_1 *in* (2.19)*, one has*

$$\Delta_1 \lesssim \begin{cases} \sigma^{-2\vartheta} h^{2\vartheta+2k} \exp\left(-2\gamma\left(\sigma/h\right)^{\beta}\right) & \text{if } h < \sigma \\ h^{2k} & \text{if } h \geq \sigma \end{cases} \tag{2.20}$$

Lemma 2.4.2. *If* $\beta > b > 0$*, denote*

$$\kappa = \left(\frac{db}{\gamma\beta}\right)^{\frac{b}{\beta-b}} \left[\frac{d(\beta-b)}{b}\right] > 0. \tag{2.21}$$

Then, under the assumptions (2.9)–(2.12)*, the expressions for* Δ_2 *defined in* (2.19)*, are given in Table 2.1.*

Observe that in every case, the expression for the variance depends not only on the values of h, σ and n but also on their mutual relationship. Also, the bias term $\Delta_1(\sigma, h)$ is an increasing function of h while the variance term $\Delta_2(\sigma, h)$ is a decreasing function of h, so the optimal value $h = h_{\text{opt}}$ is such that $\Delta_1(\sigma, h) \asymp n^{-1}\, \Delta_2(\sigma, h)$. Theorem 2.4.1 below presents the optimal expressions h_{opt} for the bandwidth h as well as the corresponding values for the risk $\Delta(n, \sigma, h_{opt})$ where $\Delta(n, \sigma, h)$ is defined in (2.17).

Theorem 2.4.1. *Let conditions* (2.9)–(2.12) *hold. Then, the asymptotic values of*

$$h_{opt} = \arg\min_h [\Delta(n, \sigma, h)]$$

Table 2.1: The asymptotic expressions for $\Delta_2 \equiv \Delta_2(\sigma, h)$

Case	Δ_2
I) $b = \beta = 0, \vartheta > a + \frac{1}{2}$,	$\min\left(h^{-(2a+1)}, \sigma^{-(2a+1)}\right)$
II) $b = \beta = 0, \vartheta = a + \frac{1}{2}$	$\min\left(h^{-(2a+1)}, \sigma^{-(2a+1)}\right) \max\left\{\ln\left(\frac{\sigma}{h}\right), 1\right\}$
III) $b = \beta = 0, \vartheta < a + \frac{1}{2}$,	$h^{-(2a+1)} \min\left\{\left(\frac{h}{\sigma}\right)^{2\vartheta}, 1\right\}$
IV) $b = 0, \beta > 0$	$\min\left(h^{-(2a+1)}, \sigma^{-(2a+1)}\right)$
V) $\beta > b > 0, h > \left(\frac{\gamma\beta}{db}\sigma^{\beta}\right)^{\frac{1}{\beta-b}}$ $\beta > b > 0, h < \left(\frac{\gamma\beta}{db}\sigma^{\beta}\right)^{\frac{1}{\beta-b}}$	$h^{-(2a+1)+b}\exp(2dh^{-b}) \min\left\{\left(\frac{h}{\sigma}\right)^{2\vartheta}, 1\right\}$ $\exp\left(\kappa\sigma^{-\frac{\beta b}{\beta-b}}\right) \sigma^{\frac{\beta}{\beta-b}\cdot\frac{b-2}{2}-2\vartheta}$
VI) $b = \beta > 0$	$h^{-(2a+1)+b} \exp(2h^{-b}(d - \gamma\sigma^{b})) \min\left\{\left(\frac{h}{\sigma}\right)^{2\vartheta}, 1\right\}$
VII) $b > 0, \beta = 0$	$h^{-(2a+1)+b}\exp(2dh^{-b}) \min\left\{\left(\frac{h}{\sigma}\right)^{2\vartheta}, 1\right\}$
VIII) $b > \beta > 0$	$h^{-(2a+1)+b}\exp(2dh^{-b}) \min\left\{\left(\frac{h}{\sigma}\right)^{2\vartheta}, 1\right\}$

and also of $\Delta(n, \sigma, h_{opt})$ *are provided in Table 2.2. Here,*

$$\begin{aligned} \mu_1 = \mu_1(n) &= \left[\frac{1}{2d}\left(\ln n + \left(\frac{b-2a-1}{b}\right)\ln\ln n\right)\right]^{-\frac{1}{b}}, \\ \mu_2 = \mu_2(n) &= \left[\frac{1}{2(d-\gamma\sigma^{b})}\left(\ln n + \left(\frac{b-2a-1}{b}\right)\ln\ln n\right)\right]^{-\frac{1}{b}}. \end{aligned} \tag{2.22}$$

Table 2.2: The optimal values of the bandwidth h and the corresponding expressions for the MISE

Case	$\Delta(n,\sigma,h_{opt})$	condition	h_{opt}
I) $b=\beta=0$, $\vartheta > a+\frac{1}{2}$	$n^{-1}\sigma^{-(2a+1)}$ $n^{-\frac{2k}{2k+2a+1}}$	$\sigma > n^{-\frac{1}{2k+2a+1}}$ $\sigma \le n^{-\frac{1}{2k+2a+1}}$	0 $n^{-\frac{1}{2k+2a+1}}$
II) $b=\beta=0$ $\vartheta = a+\frac{1}{2}$	$n^{-1}\sigma^{-(2a+1)}\ln n$ $n^{-\frac{2k}{2k+2a+1}}$	$\sigma > n^{-\frac{1}{2k+2a+1}}$ $\sigma \le n^{-\frac{1}{2k+2a+1}}$	$n^{-\frac{1}{2k+2a+1}}$ $n^{-\frac{1}{2k+2a+1}}$
III) $b=\beta=0$, $\vartheta < a+\frac{1}{2}$	$\sigma^{-2\vartheta}n^{-\frac{2\vartheta+2k}{2k+2a+1}}$ $n^{-\frac{2k}{2k+2a+1}}$	$\sigma > n^{-\frac{1}{2k+2a+1}}$ $\sigma \le n^{-\frac{1}{2k+2a+1}}$	$n^{-\frac{1}{2k+2a+1}}$ $n^{-\frac{1}{2k+2a+1}}$
IV) $b=0, \beta>0$	$n^{-1}\sigma^{-(2a+1)}$ $n^{-\frac{2k}{2k+2a+1}}$	$\sigma > n^{-\frac{1}{2k+2a+1}}$ $\sigma \le n^{-\frac{1}{2k+2a+1}}$	0 $n^{-\frac{1}{2k+2a+1}}$
V) $\beta > b > 0$	$n^{-1}\exp\left(\kappa\sigma^{\frac{-\beta b}{\beta-b}}\right)\sigma^{\frac{\beta(b-2)}{2(\beta-b)}-2\vartheta}$ $(\ln n)^{-\frac{2k}{b}}$	$\sigma > \mu_1$ $\sigma \le \mu_1$	0 μ_1
VI) $b=\beta>0$	$\sigma^{-2\vartheta}(\ln n)^{-\frac{2\vartheta+2k}{b}}\exp\left(-2\gamma\sigma^{\beta}(\ln n)^{\frac{\beta}{b}}\right)$ $(\ln n)^{-\frac{2k}{b}}$	$\sigma > \mu_1$ $\sigma \le \mu_1$	μ_1 μ_2
VII) $b>0, \beta=0$	$(\ln n)^{-\frac{(2\vartheta+2k)}{b}}\sigma^{-2\vartheta}$ $(\ln n)^{-\frac{2k}{b}}$	$\sigma > \mu_1$ $\sigma \le \mu_1$	μ_1 μ_1
VIII) $b > \beta > 0$	$\sigma^{-2\vartheta}(\ln n)^{\frac{(1+2a-2\vartheta)}{b}-1}$ $(\ln n)^{-\frac{2k}{b}}$	$\sigma > \mu_1$ $\sigma \le \mu_1$	μ_1 μ_1

The optimal values h_{opt} of the bandwidth h and the corresponding expressions for $\Delta(n,\sigma,h)$ defined in (2.18). Here, μ_1 and μ_2 are given by (2.22).

2.5 Adaptive Estimation Using Lepski's Method

Note that although Theorem 2.4.1 provides the optimal values for the bandwidth and the corresponding convergence rates, in practice, we can use those values only in the cases V-VIII, since in the cases I-IV the value of the optimal bandwidth h_{opt} depends on the smoothness parameter k of the unknown density f_X. Moreover, in cases I and IV the optimal bandwidth is zero if $\sigma > n^{-\frac{1}{2k+2a+1}}$ where the threshold value $n^{-\frac{1}{2k+2a+1}}$ itself depends on the unknown value of k. In order to resolve this difficulty, we use a novel modification of the Lepski method for construction of adaptive estimators (see, e.g., Lepski and Spokoiny (1997), Goldenshluger and Lepski (2011)).

Below we consider cases I-IV, for which the optimal value h_{opt} depends on the unknown parameter k. To start with, note that, by Lemma 2.4.1, if $h_{opt} = 0$, as it happens in cases I and IV, one has

$$\Delta(n, \sigma, 0) \asymp \Delta(n, \sigma, n^{-1})$$

Moreover, if $\sigma \leq n^{-\frac{1}{2a+1}} < n^{-\frac{1}{2k+2a+1}}$, then $h_{opt} > 1/n$.

In order to replace the unknown value of h_{opt} by its estimated value, we use the variance term given by

$$\begin{aligned} D(n, \sigma, h) &= \|\hat{f}_{W,h}(x) - f_{W,h}(x)\|^2 = \tfrac{1}{2\pi} \|\hat{f}^*{}_{W,h}(x) - f^*{}_{W,h}(x)\|^2 \\ &= \tfrac{1}{2\pi} \int_{-1/h}^{1/h} \frac{|g^*(\sigma s)|^2}{|f^*_\xi(s)|^2} \, |\hat{f}^*_Y(s) - f^*_Y(s)|^2 \, ds. \end{aligned}$$

If $h \geq 1/n$, then it is easy to see that

$$D(n, \sigma, h) \leq \max_{|s| \leq n} |\hat{f}^*_Y(s) - f^*_Y(s)|^2 \; I(\sigma, h)$$

where $I(\sigma, h)$ is defined in (2.16).

Recall also that the value h_{opt} is such that it minimizes the sum of $\Delta_1(n, \sigma, h) + n^{-1}\Delta_2(n, \sigma, h)$ where, under the assumptions A1 - A3, the first term is growing polynomially in h while the second is decreasing polynomially in h. Therefore,

$$\Delta(n, \sigma, h_{opt}) \asymp \Delta_1(n, \sigma, h_{opt}) \asymp n^{-1} \, I(n, \sigma, h_{opt}). \tag{2.23}$$

Consider the sets

$$\mathcal{J} = \{1, 2, 3, \cdots, j_{\max}\} \quad \text{and} \quad \mathcal{H} = \{h = 2^{-j}, j \in \mathcal{J}\} \tag{2.24}$$

and denote

$$j_{\max} = \min\left(\frac{\ln n}{2a+1}, \ln\left(\frac{1}{\sigma}\right)\right). \tag{2.25}$$

Let $q > 0$ be such that $\mathbb{E}(|X_1|^q) \le C_q < \infty$ and

$$C(\tau, q) \ge 8\sqrt{2\tau(q+1) + 6q + 2}/\sqrt{q}. \tag{2.26}$$

Define a set in the sample space

$$\Omega_{\sigma,n} = \begin{cases} \{w : \| \hat{f}_{W,\sigma} - f_{W,\frac{1}{n}} \| \ge 4C(\tau,q)\sqrt{n^{-1}\, I(\sigma, 1/n) \ln n}\} & \text{for cases I and IV} \\ \emptyset \text{ (the empty set)} & \text{for cases II and III.} \end{cases} \tag{2.27}$$

Then, the following statement holds.

Theorem 2.5.1. *Let conditions* (2.9)–(2.12) *hold with* $b = 0$ *(cases I-IV) and* $\tau \ge 4$. *Define*

$$\hat{h} = \begin{cases} 1/n & \textit{if } w \in \Omega_{\sigma,n} \\ \max\{h \in \mathcal{H} : \|\hat{f}_{W,h} - f_{W,\tilde{h}}\| \le 4C(\tau,q)\sqrt{\frac{\ln n}{n} I(\sigma, \tilde{h})} \;\; \textit{for any} \;\; \tilde{h} \le h, \tilde{h} \in \mathcal{H}\} & \textit{if } w \notin \Omega_{\sigma,n} \end{cases}$$

Then

$$\mathbb{E}\|\hat{f}_{W,\hat{h}} - f_W\|^2 \lesssim \Delta(n, \sigma, h_{opt}) \ln n. \tag{2.28}$$

2.6 Discussion

In this work, our main goal was to justify the choice of a bandwidth in deconvolution problems with small Berkson errors. To the best of our knowledge, our paper is the first paper which carries out a comprehensive theoretical study of density deconvolution with Berkson errors when Berkson errors are asymptotically small.

In particular, we refined the conclusion of Long *et al.* (2016) and studied the relationship between the three parameters: the bandwidth h, the sample size n and the standard deviation of the Berkson errors σ. As Theorem 2.4.1 above shows, the expressions for the optimal bandwidth are always chosen to minimize the error in the estimator of the density of interest f_W. In particular, if $h = 0$ is possible, one should choose this value as long as the Berkson errors are not too small, i.e., σ lies above some threshold level that depends on the shapes of the densities and the number of observations n.

In order to uncover the reason for this, compare expressions (2.6) and (2.8) and observe that $g^*(\sigma s)$ in (2.6) acts as a kernel function g with the bandwidth $h = \sigma$. If σ is large enough (i.e $\sigma > h_{opt}$, where h_{opt} is the value of h that achieves the best bias-variance balance), then convolution with g leads to sufficient regularization and no kernel estimation is necessary. However , if $\sigma < h_{opt}$ then one needs additional kernel smoothing with $h > \sigma$.

The setting of Long *et al.* (2016) corresponds to the cases I, II, III and IV in Tables 1 and 2 with $a = b = 0$. If $\vartheta > 1/2$, then h_{opt} is zero if σ is large enough and h_{opt} is of the order $n^{-1/(2k+1)}$ (where k is the degree of smoothness of the density f_X of the measurements) otherwise. The choice depends on the relationship between parameters σ, n and k. Since k is unknown, we construct adaptive estimators of f_W using a novel modification of Lepski method. Indeed, one cannot use the traditional Lepski method since the value of the optimal bandwidth depends on the relationship between σ and the unknown threshold $n^{-1/(2k+2a+1)}$. Hence, our paper presents a non-trivial extension of the Lepski technique.

Note that we did not consider the case of the multivariate density functions. This extension is fairly straightforward but rather cumbersome. We shall leave this case for the future investigation.

2.7 Supplementary Lemmas and Proofs

2.7.1 Proof of Lemma 2.4.1.

Proof. Since for any $f_X \in S(k, B)$ one has

$$
\begin{aligned}
\Delta_1 &= \max_{f_X \in \mathcal{S}(k,B)} \|\mathbb{E}\hat{f}_{W,h} - f_W\|^2. \\
&= \max_{f_X \in \mathcal{S}(k,B)} \frac{1}{2\pi} \int_{|s|>1/h} |g^*(\sigma s)|^2 |f_X^*(s)|^2 ds \\
&= \max_{f_X \in \mathcal{S}(k,B)} \frac{1}{\pi} \int_{\frac{1}{h}}^{\infty} |g^*(\sigma s)|^2 |f_X^*(s)|^2 ds \\
&\le \max_{f_X \in \mathcal{S}(k,B)} \frac{2C_g}{\pi} \int_{\frac{1}{h}}^{\infty} (\sigma^2 s^2 + 1)^{-\vartheta} \exp(-2\gamma |s|^\beta \sigma^\beta) \frac{(s^2+1)^k}{(s^2+1)^k} |f_X^*(s)|^2 ds \\
&\le \frac{2C_g B^2}{\pi} \max_{s \ge \frac{1}{h}} \left[(\sigma^2 s^2 + 1)^{-\vartheta} \exp(-2\gamma |s|^\beta \sigma^\beta) \right] (h^{-2} + 1)^{-k},
\end{aligned}
$$

obtain

$$
\Delta_1 \lesssim \min\left\{ \left(\frac{h}{\sigma}\right)^{2\vartheta}, 1 \right\} h^{2k} \exp\left(-2\gamma \left(\frac{\sigma}{h}\right)^\beta \right)
$$

which implies (2.20). □

2.7.2 *Proof of Lemma 2.4.2*

Proof. Note that the variance term is given by

$$
\begin{aligned}
\Delta_2 &\leq \frac{1}{2\,\pi}\int_{-\infty}^{\infty}\frac{|g^*(\sigma s)|^2}{|f^*_\xi(s)|^2}I(|s|<h^{-1})ds \\
&\leq \frac{C_g}{c_\xi}\int_0^{\frac{1}{h}}(\sigma^2 s^2+1)^{-\vartheta}\,(s^2+1)^a\,\exp(-2\gamma|s|^\beta\sigma^\beta+2d|s|^b)\,ds
\end{aligned}
$$

Using change of variables s=z/h obtain

$$
\Delta_2 \lesssim h^{-(2a+1)}\,V(\sigma,h) \quad \text{with} \quad V(\sigma,h)=\int_0^1 P(z|\sigma,h)\exp\{\phi(z|\sigma,h)\}dz \tag{2.29}
$$

where

$$
\phi(z|\sigma,h)=2dz^b h^{-b}-2\gamma z^\beta\sigma^\beta h^{-\beta}, \quad P(z|\sigma,h)=\left(\sigma^2 z^2 h^{-2}+1\right)^{-\vartheta}\left(z^2+h^2\right)^a \tag{2.30}
$$

For cases when $b=0$ (cases I-IV), one can obtain an asymptotic expression for Δ_2 using direct calculations. If $b>0$ and $d\geq 0$, one needs to apply Lemma 2.7.4. Denote by z_0 and z_h, respectively, the point where $\phi(z|\sigma,h)$ attains its global maximum on the interval $[0,1]$ and its critical point:

$$
z_0\equiv z_0(\sigma,h)=\operatorname*{argmax}_{z\in[0,1]}\phi(z|\sigma,h), \quad z_h=\left(db\,(\gamma\beta)^{-1}\,\sigma^{-\beta}\right)^{\frac{1}{\beta-b}}\,h \tag{2.31}
$$

Since $z_h>0$, there are two possible cases here: $z_h\in(0,1]$ and $z_h>1$. If $z_h\in(0,1]$, then $z_0=z_h$, $\phi'(z_0)=0$ and $\phi''(z_0)<0$. If $z_h>1$, then $z_0=1$ and $\phi'(z_0)=\phi'(1)>0$.

Hence, Lemma 2.7.4 and formula (2.29) yield, that for small values of h and σ,

$$h^{2a+1}\,\Delta_2 \lesssim \begin{cases} \dfrac{\exp\{\phi(z_h|\sigma,h)\}P(z_h|\sigma,h)}{\sqrt{|\phi''(z_h|\sigma,h)|}}, & \text{if} \quad z_0 = z_h, \\ \dfrac{\exp\{\phi(1|\sigma,h)\}P(1|\sigma,h)}{\phi'(1|\sigma,h)}, & \text{if} \quad z_0 = 1. \end{cases} \tag{2.32}$$

Here

$$\begin{aligned} &\phi(1|\sigma,h) = 2dh^{-b} - 2\gamma\sigma^{\beta}h^{-\beta}, \quad \phi'(1|\sigma,h) = 2\left(dbh^{-b} - \gamma\beta\sigma^{\beta}h^{-\beta}\right) \\ &P(1|\sigma,h) \asymp \left(\sigma^2 h^{-2} + 1\right)^{-\vartheta}, \qquad P(z_h|\sigma,h) = \left(\sigma^2 z_h^2 h^{-2} + 1\right)^{-\vartheta}\left(z_h^2 + h^2\right)^a \end{aligned} \tag{2.33}$$

Below we consider various cases.

Cases I, II, III: $b = \beta = 0$.

Note that

$$I(\sigma,h) = \frac{2{C_g}^2}{{c_\xi}^2}\int_0^{\frac{1}{h}}(\sigma^2 s^2+1)^{-\vartheta}(s^2+1)^a ds = \frac{2{C_g}^2}{{c_\xi}^2 h}\int_0^1 \left(\sigma^2 z^2 h^{-2}+1\right)^{-\vartheta}\,\left(z^2h^{-2}+1\right)^a\,dz \tag{2.34}$$

If $h \geq \sigma$, then $\sigma^2 z^2 h^{-2} + 1 \in (1,2)$ and $I(\sigma,h) \leq \dfrac{2^{1-\vartheta}{C_g}^2}{{c_\xi}^2}h^{-(2a+1)}$.

If $h < \sigma$, then, by the change of variables $\sigma s = u$ in (2.34), obtain

$$I(\sigma,h) = \frac{2{C_g}^2}{{c_\xi}^2\sigma}\int_0^{\frac{\sigma}{h}}\left(u^2+1\right)^{-\vartheta}\,\left(u^2\sigma^{-2}+1\right)^a\,du \leq \frac{2{C_g}^2 C_a}{{c_\xi}^2}\sigma^{-(2a+1)}\int_0^{\frac{\sigma}{h}}\frac{u^{2a}}{(u^2+1)^{\vartheta}}du$$

Hence,

$$I(\sigma,h) \leq \frac{2{C_g}^2 C_a}{{c_\xi}^2}\min\left(h^{-(2a+1)},\sigma^{-(2a+1)}\right)\Delta_{h\sigma}$$

where

$$\Delta_{h,\sigma} = \begin{cases} 1 & \text{if } \vartheta > a + 1/2 \\ \max\left\{\ln\left(\frac{\sigma}{h}\right), 1\right\} & \text{if } \vartheta = a + 1/2 \\ \max\left\{1, \left(\frac{\sigma}{h}\right)^{2a-2\vartheta+1}\right\} & \text{if } \vartheta < a + 1/2 \end{cases} \tag{2.35}$$

Case IV: $b = 0, \beta > 0$.

In this case,

$$\Delta_2 \asymp h^{-1} \int_0^1 \left(\sigma^2 z^2 h^{-2} + 1\right)^{-\vartheta} \left(z^2 h^{-2} + 1\right)^a \exp\left(-2\gamma\sigma^\beta z^\beta h^{-\beta}\right) dz.$$

If $h > \sigma$ then the argument of the exponent is bounded above and $\Delta_2 \asymp h^{-2a-1}$. If $h < \sigma$, then by changing variables $u = 2\gamma\left(\sigma z/h\right)^\beta$, obtain

$$\Delta_2 \asymp \sigma^{-1} \int_0^\infty \left(\left(\frac{u}{2\gamma}\right)^{\frac{2}{\beta}} + 1\right)^{-\vartheta} \left(\frac{1}{\sigma^{2a}}\left(\frac{u}{2\gamma}\right)^{\frac{2a}{\beta}} + 1\right) \exp(-u) u^{\frac{1}{\beta}-1} du \asymp \sigma^{-(2a+1)}$$

Hence,

$$\Delta_2 \asymp \min\left(h^{-(2a+1)}, \sigma^{-(2a+1)}\right).$$

Case V: $\beta > b > 0$.

In this case $\rho^2(\sigma) = \infty$ in (2.5), so that $h > 0$. The expression for the variance is given by (2.29) with $\phi(z|\sigma, h)$ defined in (2.30). Let z_h be given by (2.31). It is easy to check that

$$z_h = \left(db\,(\gamma\beta)^{-1}\sigma^{-\beta}\right)^{\frac{1}{\beta-b}} h \asymp \sigma^{-\frac{\beta}{\beta-b}} h. \tag{2.36}$$

It is easy to check that $\phi''(z_h|\sigma,h) < 0$, so that z_h is the local maximum. Now consider two cases.

(a) If $h > \left(\frac{\gamma\beta}{db}\sigma^\beta\right)^{\frac{1}{\beta-b}}$, then $z_h > 1$. Hence, $\phi(z|\sigma,h)$ does not have a local maximum on $[0,1]$ and it attains its global maximum at $z_0 = 1$. Then, $2dh^{-b} > \phi(1|\sigma,h) = 2dh^{-b} - 2\gamma\sigma^\beta h^{-\beta} > 2dh^{-b}\left(1 - b/\beta\right)$. Moreover, since $\beta > b$ and $h > \left(\frac{\gamma\beta}{db}\sigma^\beta\right)^{\frac{1}{\beta-b}} > \sigma$, one has $2dbh^{-b} > 2\gamma\beta\sigma^\beta h^{-\beta}$ which yields

$$\phi'(1|\sigma,h) = 2dbh^{-b} - 2\gamma\beta\sigma^\beta h^{-\beta} = 2dbh^{-b}\left(1 - \frac{\gamma\beta}{db}\sigma^\beta h^{b-\beta}\right) \asymp h^{-b}$$

Plugging those expressions into the second equation of (2.32) and using (2.33), obtain

$$\Delta_2 \asymp h^{-(2a+1)} \min\left\{\left(h\sigma^{-1}\right)^{2\vartheta}, 1\right\} \exp(2dbh^{-b})\, h^b \asymp h^{b-2a-1} \exp\left(2dh^{-b}\right)$$

(b) If $h < \left(\frac{\gamma\beta}{db}\sigma^\beta\right)^{\frac{1}{\beta-b}}$, then z_h is given by formula (2.36) and $z_0 = z_h < 1$. Hence, Δ_2 is given by the first expression in formula (2.32)

$$\Delta_2 \asymp \frac{\exp(\phi(z_h|\sigma,h))}{\sqrt{|\phi''(z_h|\sigma,h)|}}\, h^{-(2a+1)} \left(\sigma^2 z_h^2 h^{-2} + 1\right)^{-\vartheta} \left(z_h^2 + h^2\right)^a \tag{2.37}$$

Note that, due to $\beta > b > 0$, $\frac{\beta^2}{\beta-b} > \frac{\beta b}{\beta-b}$ and $\beta - \frac{\beta^2}{\beta-b} = -\frac{\beta b}{\beta-b}$, one has

$$\phi(z_h|\sigma,h) = \frac{2d}{h^b}\left(\frac{db}{\gamma\beta}\sigma^{-\beta}\right)^{\frac{b}{\beta-b}} h^b - \frac{2\gamma\sigma^\beta}{h^\beta}\left(\frac{db}{\gamma\beta}\sigma^{-\beta}\right)^{\frac{\beta}{\beta-b}} h^\beta = \kappa\sigma^{-\frac{\beta b}{\beta-b}}$$

where κ is a positive constant defined in (2.21). Also

$$\phi''(z_h|\sigma,h) = \frac{2}{z_h^2}\left(\frac{db(b-1)z_h^b}{h^b} - \frac{\gamma\beta(\beta-1)\, z_h^\beta\sigma^\beta}{h^\beta}\right) = \frac{2db(b-\beta)z_h^{b-2}}{h^b} \asymp \frac{z_h^{b-2}}{h^b}$$

Then, plugging $\phi(z_h|\sigma, h)$ and $\phi''(z_h|\sigma, h)$ into (2.37), obtain

$$\Delta_2 \asymp \exp\left(\kappa\sigma^{-\frac{\beta b}{\beta - b}}\right)\sigma^{\frac{\beta(b-2)}{2(\beta-b)}-2\vartheta}.$$

Case VI: $b = \beta > 0$.

In this case $\rho^2(\sigma) = \infty$ in (2.5), so that $h > 0$. Moreover, since $\phi(z|\sigma, h) = 2z^b\, h^{-b}(d - \gamma\sigma^b)$ where, due to condition (2.12), $d - \gamma\sigma^b > 0$, $z_0 = 1$ is the non-local maximum of $\phi(z|\sigma, h)$. Then, the second expression in formula (2.32)

$$\Delta_2 \lesssim \frac{\exp(\phi(1|\sigma, h))}{\phi'(1|\sigma, h)}\, h^{-(2a+1)}\left(\sigma^2 h^{-2} + 1\right)^{-\vartheta} \tag{2.38}$$

Using (2.33) with $\beta = b$, we derive

$$\Delta_2 \lesssim h^{b-(2a+1)} \min\left(\left(\frac{h}{\sigma}\right)^{2\vartheta}, 1\right) \exp(2h^{-b}(d - \gamma\sigma^b))$$

Case VII: $b > 0, \beta = \gamma = 0$

In this case, $z_0 = 1$ is the non-local maximum of $\phi(z|\sigma, h)$ and (2.33) yield $\phi(1|\sigma, h) = 2dh^{-b}$ and $\phi'(1|\sigma, h) = 2dbh^{-b}$. Plugging those expressions into (2.38), we derive

$$\Delta_2 \lesssim \min\left(\left(\frac{h}{\sigma}\right)^{2\vartheta}, 1\right) h^{b-(2a+1)}\exp(2dh^{-b})$$

Case VIII: $b > \beta > 0$

In this case $\rho^2(\sigma) = \infty$ in (2.5), so that $h > 0$. Also, it is easy to check that although $z_h \in (0, 1)$, one has $\phi''(z_h|\sigma, h) > 0$, so z_h is the local minimum. It is easy to see that $z_0 = 1$ and $\phi(1|\sigma, h) = 2dh^{-b}(1 - \gamma d^{-1}\sigma^\beta h^{b-\beta}) \asymp 2dh^{-b}$. Moreover, $\phi'(1|\sigma, h) = 2h^{-b}(db - \gamma\beta\sigma^\beta h^{b-\beta}) \asymp h^{-b}$, so formula (2.38) yields

$$\Delta_2 \lesssim h^{b-(2a+1)} \min\left(\left(\frac{h}{\sigma}\right)^{2\vartheta}, 1\right) \exp(2dh^{-b}).$$

□

2.7.3 Proof of Theorem 2.4.1

Proof. Consider various cases.

Cases I, II, III: $b = \beta = 0$.

One has

$$\Delta \lesssim \min\left\{\left(h\,\sigma^{-1}\right)^{2\vartheta}, 1\right\} h^{2k} + n^{-1} \min\left(h^{-(2a+1)}, \sigma^{-(2a+1)}\right) \Delta_{h\sigma} \tag{2.39}$$

where $\Delta_{h,\sigma}$ is defined in (2.35).

Case I: $b = \beta = 0, \vartheta > a + 1/2$.

In this case $\rho^2(\sigma) < \infty$ and $h = 0$ is possible. If $h = 0$, then $\Delta = O\left(\sigma^{-(2a+1)} n^{-1}\right)$. If $h \neq 0$, then choose $h \geq \sigma$, so that $\Delta_1(\sigma, h) \lesssim h^{2k}$, $\Delta_2(\sigma, h) \lesssim h^{-(2a+1)}$. Then, $h_{opt} \asymp n^{-\frac{1}{2k+2a+1}}$ and $\Delta_1(\sigma, h_{opt}) + n^{-1}\,\Delta_2(\sigma, h_{opt}) \lesssim n^{-\frac{2k}{2k+2a+1}}$. Choose $h = h_{opt}$ if $h_{opt} \geq \sigma$, i.e., if $n^{-\frac{1}{2k+2a+1}} \geq \sigma$.

Obtain

$$\Delta \asymp \begin{cases} n^{-1}\,\sigma^{-(2a+1)}, & h_{opt} = 0 \qquad\qquad \text{if } \sigma > n^{-\frac{1}{2k+2a+1}} \\ n^{-\frac{2k}{2k+2a+1}}, & h_{opt} = n^{-\frac{1}{2k+2a+1}} \quad \text{if } \sigma \le n^{-\frac{1}{2k+2a+1}} \end{cases}$$

Case II: $b = \beta = 0$, $\vartheta = a + \frac{1}{2}$.

Here, Δ is given by (2.39) where $\Delta_{h\sigma} = \max\left\{\ln(\sigma/h), 1\right\}$. If $h < \sigma$, then $\Delta \lesssim \sigma^{-2\vartheta} h^{2\vartheta+2k} + \sigma^{-(2a+1)} n^{-1} \ln(\sigma/h)$. Setting $\sigma^{-2\vartheta} h^{2\vartheta+2k} = \sigma^{-(2a+1)} n^{-1} \ln(\sigma/h)$ leads to

$$h_{opt} \asymp n^{-\frac{1}{2k+2a+1}}, \quad \Delta \lesssim n^{-1}\,\sigma^{-(2a+1)}\,\ln n.$$

Note that $h_{opt} < \sigma$ if and only if $n^{-\frac{1}{2k+2a+1}} < \sigma$. Now, consider the case when $h \ge \sigma$. Then by (2.39), $\Delta \lesssim n^{-\frac{2k}{2k+2a+1}}$ if $n^{-\frac{1}{2k+2a+1}} \ge \sigma$. Hence

$$\Delta \asymp \begin{cases} \dfrac{\sigma^{-(2a+1)}}{n} \ln n, & h_{opt} = n^{-\frac{1}{2k+2a+1}} \quad \text{if } \sigma > n^{-\frac{1}{2k+2a+1}} \\ n^{-\frac{2k}{2k+2a+1}}, & h_{opt} \asymp n^{-\frac{1}{2k+2a+1}} \quad \text{if } \sigma \le n^{-\frac{1}{2k+2a+1}} \end{cases}$$

Case III: $b = \beta = 0$, $\vartheta < a + \frac{1}{2}$.

First, consider the case when $h < \sigma$. Then, by (2.39) and (2.35), obtain

$$\Delta \lesssim \sigma^{-2\vartheta} h^{2\vartheta+2k} + \sigma^{-(2\vartheta)} n^{-1} h^{2\vartheta-2a-1}.$$

Setting $\sigma^{-2\vartheta} h^{2\vartheta+2k} = \sigma^{-(2\vartheta)} n^{-1} h^{2\vartheta-2a-1}$, obtain $h_{opt} \asymp n^{-\frac{1}{2k+2a+1}}$ and $\Delta \lesssim \sigma^{-2\vartheta} n^{-\frac{2\vartheta+2k}{2k+2a+1}}$. Also note that $h_{opt} < \sigma$ if and only if $\sigma > n^{-\frac{1}{2k+2a+1}}$. Now, consider the case when $h \ge \sigma$. Then, (2.39) and (2.35), derive that $\Delta \asymp n^{-\frac{2k}{2k+2a+1}}$ if $n^{-\frac{1}{2k+2a+1}} \ge \sigma$. Hence

$$\Delta \asymp \begin{cases} \sigma^{-2\vartheta} n^{-\frac{2\vartheta+2k}{2k+2a+1}}, & h_{opt} = n^{-\frac{1}{2k+2a+1}} \quad \text{if } \sigma > n^{-\frac{1}{2k+2a+1}} \\ n^{-\frac{2k}{2k+2a+1}}, & h_{opt} = n^{-\frac{1}{2k+2a+1}} \quad \text{if } \sigma \le n^{-\frac{1}{2k+2a+1}} \end{cases}$$

Case IV: $b = 0, \beta > 0$.

In this case $\rho^2(\sigma) < \infty$ and $h = 0$ is possible. Consider the case $h < \sigma$. Then,

$$\Delta_1(\sigma, h) \lesssim \sigma^{(-2\vartheta)} h^{2\vartheta+2k} \exp\left(-2\gamma \left(\frac{\sigma}{h}\right)^{\beta}\right), \quad \Delta_2(\sigma, h) \lesssim \sigma^{-(2a+1)}.$$

If $h < \sigma$, then $h_{opt} = 0$ and $\Delta \asymp n^{-1}\sigma^{-(2a+1)}$. If $h > \sigma$, then $\Delta_1(\sigma, h) \leq h^{2k}$ and $\Delta_2(\sigma, h) \lesssim h^{-(2a+1)}$. Therefore, $h_{opt} \asymp n^{-\frac{1}{2k+2a+1}}$ and $\Delta \lesssim n^{\frac{-2k}{2k+2a+1}}$. Observing that $h_{opt} \geq \sigma$ if $\sigma \leq n^{-\frac{1}{2k+2a+1}}$, obtain

$$\Delta \asymp \begin{cases} n^{-1}\sigma^{-(2a+1)} & h_{opt} = 0 & \text{if } \sigma > n^{-\frac{1}{2k+2a+1}} \\ n^{-\frac{2k}{2k+2a+1}} & h_{opt} = n^{-\frac{1}{2k+2a+1}} & \text{if } \sigma \leq n^{-\frac{1}{2k+2a+1}} \end{cases}$$

Case V: $\beta > b > 0$.

In this case $\rho^2(\sigma) < \infty$ and $h = 0$ is possible. The bias is given by (2.20) and

$$\Delta_2 \lesssim \begin{cases} n^{-1}h^{b-2a-1}\exp\left(2dh^{-b}\right) & \text{if } h > \left(\frac{\gamma\beta}{db}\sigma^{\beta}\right)^{\frac{1}{\beta-b}} \\ n^{-1}\exp\left(\kappa\sigma^{\frac{-\beta b}{\beta-b}}\right)\sigma^{\frac{\beta}{\beta-b}\cdot\frac{b-2}{2}-2\vartheta} & \text{if } h < \left(\frac{\gamma\beta}{db}\sigma^{\beta}\right)^{\frac{1}{\beta-b}} \end{cases}$$

If $h = 0$, then $\Delta \asymp n^{-1}\exp\left(\kappa\sigma^{\frac{-\beta b}{\beta-b}}\right)\sigma^{\frac{\beta}{\beta-b}\cdot\frac{b-2}{2}-2\vartheta}$. If $h > 0$, then one needs $h > \sigma \gtrsim \left(\frac{\gamma\beta}{db}\sigma^{\beta}\right)^{\frac{1}{\beta-b}}$ and $\Delta \asymp h^{2k} + n^{-1}h^{b-2a-1}\exp\left(2dh^{-b}\right)$. Choosing h such that $h^{2k} = n^{-1}h^{b-2a-1}\exp\left(2dh^{-b}\right)$, arrive at

$$(2dh^{-b})^{\frac{2a+2k+1-b}{b}}\exp\left(2dh^{-b}\right) = (2d)^{\frac{2a+2k+1-b}{b}}\, n \tag{2.40}$$

and, by Lemma 2.7.5, obtain $h_{opt} = \mu_1(n)$ where $\mu_1(n)$ is defined in (2.22), and, hence, $\Delta \asymp$

$(\ln n)^{-\frac{2k}{b}}$. Therefore,

$$\Delta \asymp \begin{cases} n^{-1} \exp\left(\kappa\sigma^{\frac{-\beta b}{\beta - b}}\right) \sigma^{\frac{\beta(b-2)}{2(\beta-b)} - 2\vartheta}, & h_{opt} = 0 \qquad \text{if } \sigma > \mu_1(n) \\ (\ln n)^{-\frac{2k}{b}}, & h_{opt} = \mu_1(n), \quad \text{if } \sigma \le \mu_1(n) \end{cases}$$

where $\mu_1(n)$ is given by (2.22).

Case VI: $b = \beta > 0, h > 0$

Note that, due to (2.12), one has $\sigma < (d\gamma^{-1})^{\frac{1}{b}}$. Consider two cases. If $h < \sigma$, then

$$\Delta_1(\sigma, h) \lesssim \sigma^{-2\vartheta} h^{2\vartheta+2k} \exp\left(-2\gamma\left(\sigma/h\right)^{\beta}\right), \quad \Delta_2(\sigma, h) \lesssim h^{(b+2\vartheta-2a-1)}\sigma^{-2\vartheta}\exp(2h^{-b}(d - \gamma\sigma^b)).$$

Then the bias-variance balance is achieved when

$$h^{(b-2k-2a-1)} \exp(2h^{-b}(d - \gamma\sigma^b) + 2\gamma\sigma^b h^{-b}) = n$$

which leads to (2.40) and, hence, $h_{opt} = \mu_1(n)$ where $\mu_1(n)$ is defined in (2.22). Therefore, $h_{opt} \asymp (\ln n)^{-\frac{1}{b}}$ and hence

$$\Delta \lesssim \sigma^{-2\vartheta} (\ln n)^{-\frac{2\vartheta+2k}{b}} \exp\left(-2\gamma\sigma^{\beta} (\ln n)^{\frac{\beta}{b}}\right).$$

If $h \ge \sigma$, then $\Delta \lesssim h^{2k} + n^{-1}h^{b-(2a+1)}\exp(2h^{-b}(d - \gamma\sigma^b))$ and the bias-variance balance is achieved when $h^{2k} \asymp n^{-1}h^{b-(2a+1)}\exp(2h^{-b}(d - \gamma\sigma^b))$. Then, by Lemma 2.7.5, we derive that $h_{opt} = \mu_2(n)$ where $\mu_2(n)$ is defined in (2.22), and $\Delta \lesssim (\ln n)^{-\frac{2k}{b}}$. Hence

$$\Delta \lesssim \begin{cases} \sigma^{(-2\vartheta)} (\ln n)^{-\frac{2\vartheta+2k}{b}} \exp\left(-2\gamma\sigma^{\beta} (\ln n)^{\frac{\beta}{b}}\right), & h_{opt} = \mu_1(n), \quad \text{if } \sigma > \mu_1(n) \\ (\ln n)^{-\frac{2k}{b}}, & h_{opt} = \mu_2(n), \quad \text{if } \sigma \le \mu_1(n) \end{cases}$$

where $\mu_1(n)$ and $\mu_2(n)$ are given by (2.22).

Case VII: $b > 0, \beta = 0$

If $h < \sigma$, then

$$\Delta \lesssim \sigma^{-2\vartheta} h^{2\vartheta+2k} + n^{-1}\, \sigma^{-2\vartheta}\, h^{2\vartheta-2a+b-1} \exp(2dh^{-b})$$

Setting $\sigma^{-2\vartheta} h^{2\vartheta+2k} = n^{-1}\, \sigma^{-2\vartheta}\, h^{2\vartheta-2a+b-1} \exp(2dh^{-b})$, arrive at (2.40) and $h_{opt} = \mu_1(n)$ where $\mu_1(n)$is defined in (2.22). Hence, $h_{opt} \asymp (\ln n)^{-1/b}$ and $\Delta \lesssim (\ln n)^{-\frac{2\vartheta+2k}{b}}\, \sigma^{-2\vartheta}$, provided $\sigma > \mu_1(n)$.

If $h \geq \sigma$, then

$$\Delta \lesssim h^{2k} + n^{-1} h^{b-2a-1} \exp(2dh^{-b}). \tag{2.41}$$

Setting $h^{2k} \approx n^{-1} h^{b-2a-1} \exp(2dh^{-b})$, arrive at (2.40), so that $h_{opt} = \mu_1(n) \asymp (\ln n)^{-1/b}$ and $\Delta \lesssim (\ln n)^{-2k/b}$ if $\sigma \leq \mu_1(n)$. Hence

$$\Delta \asymp \begin{cases} (\ln n)^{-\frac{2\vartheta+2k}{b}}\, \sigma^{-2\vartheta}, & h_{opt} = \mu_1(n), \quad \text{if } \sigma > \mu_1(n) \\ (\ln n)^{-\frac{2k}{b}}, & h_{opt} = \mu_1(n), \quad \text{if } \sigma \leq \mu_1(n), \end{cases}$$

where $\mu_1(n)$is defined in (2.22).

Case VIII: $b > \beta > 0$

If $h \leq \sigma$, then

$$\Delta(\sigma, h) \lesssim \sigma^{-2\vartheta} h^{2\vartheta+2k} \exp\left(-2\gamma\sigma^{\beta} h^{-\beta}\right) + n^{-1} h^{2\vartheta+b-(2a+1)} \sigma^{-2\vartheta} \exp(2dh^{-b}).$$

Then, the minimum of $\Delta(\sigma, h)$ is attained if $n \asymp h^{b-(2a+1)-2k} \exp(2dh^{-b} + 2\gamma\sigma^{\beta} h^{-\beta})$. Note that, due to $\sigma^\beta < (d/\gamma)\, h^{-(b-\beta)}$, $b > \beta$ and $\sigma < 1$, one has $2dh^{-b} > 2\gamma\sigma^\beta h^{-\beta}$. Therefore, we arrive at (2.40), so that $h_{opt} \asymp (\ln n)^{-1/b}$ and $\Delta \lesssim \sigma^{-2\vartheta}(\ln n)^{\frac{(1+2a-2\vartheta)}{b}-1}$.

If $h > \sigma$, then $\Delta \lesssim h^{2k} + n^{-1}h^{b-(2a+1)}\exp(2dh^{-b})$ which coincides with (2.41) and we obtain the same expressions for h_{opt} and Δ as in that case. Hence

$$\Delta \asymp \begin{cases} \sigma^{-2\vartheta}\,(\ln n)^{\frac{(1+2a-2\vartheta)}{b}-1}, & h_{opt} = \mu_1(n), \quad \text{if } \sigma > \mu_1(n) \\ (\ln n)^{-\frac{2k}{b}}, & h_{opt} = \mu_1(n), \quad \text{if } \sigma \le \mu_1(n), \end{cases}$$

where $\mu_1(n)$is defined in (2.22). □

2.7.4 *Proof of Theorem 2.5.1*

Proof. Observe that

$$\mathbb{E}\|\hat{f}_{W,\hat{h}} - f_W\|^2 = \widetilde{\Delta}_1 + \widetilde{\Delta}_2 + \widetilde{\Delta}_3 \tag{2.42}$$

where

$$\begin{aligned} \widetilde{\Delta}_1 &= \mathbb{E}\left[\|\hat{f}_{W,\frac{1}{n}} - f_W\|^2 I(w \in \Omega_{\sigma,n})\right]\, I\left(\sigma > n^{-\frac{1}{2a+1}}\right) \\ \widetilde{\Delta}_2 &= \sum_{j=1}^{j_{opt}} \mathbb{E}\left[\|\hat{f}_{W,h} - f_W\|^2 I(\hat{h} = h = 2^{-j}) I(w \notin \Omega_{\sigma,n} \text{ or } \sigma \le n^{-\frac{1}{2a+1}})\right] \\ \widetilde{\Delta}_3 &= \sum_{j=j_{opt}+1}^{j_{\max}} \mathbb{E}\|\hat{f}_{W,h} - f_W\|^2\, I(\hat{h} = h = 2^{-j})\, I(w \notin \Omega_{\sigma,n} \text{ or } \sigma \le n^{-\frac{1}{2a+1}}) \end{aligned}$$

We start with construction of an upper bound for $\widetilde{\Delta}_1$. Consider the cases I and IV, since, otherwise, $\widetilde{\Delta}_1 = 0$. Then

$$\begin{aligned} \widetilde{\Delta}_1 &= \mathbb{E}\left[\|\hat{f}_{W,\frac{1}{n}} - f_W\|^2 I(w \in \Omega_{\sigma,n})\right]\, I\left(\sigma > n^{-\frac{1}{2a+2k+1}} = h_{opt}\right) \\ &+ \mathbb{E}\left[\|\hat{f}_{W,\frac{1}{n}} - f_W\|^2 I(w \in \Omega_{\sigma,n})\right]\, I\left(n^{-\frac{1}{2a+1}} < \sigma \le n^{-\frac{1}{2a+2k+1}}\right) = \widetilde{\Delta}_{11} + \widetilde{\Delta}_{12}. \end{aligned}$$

Here

$$\begin{aligned}\widetilde{\Delta}_{11} &\leq \mathbb{E}\|\hat{f}_{W,\frac{1}{n}} - f_W\|^2\, I(\sigma > h_{opt}) \\ &\leq C\left[\sigma^{-2\vartheta} n^{-(2\vartheta+2k)} + n^{-1}\sigma^{-(2a+1)}\right] \\ &\leq C n^{-1}\sigma^{-(2a+1)} = C\Delta_{opt} \equiv C\Delta(n,\sigma,h_{opt})\end{aligned}$$

For $\widetilde{\Delta}_{12}$, one has

$$\widetilde{\Delta}_{12} \leq \sqrt{\mathbb{E}\|\hat{f}_{W,\frac{1}{n}} - f_W\|^4}\,\sqrt{\mathbb{P}\left[(w \in \Omega_{\sigma,n})\; I\left(\sigma \leq n^{-\frac{1}{2a+2k+1}}\right)\right]}$$

By Lemma 2.7.2, in cases I and IV, $\mathbb{E}\|\hat{f}_{W,\frac{1}{n}} - f_W\|^4 \leq Cn^2$ and $\widetilde{\Delta}_{12} \leq C[n\, n^{-\frac{\tau}{2}}] \leq Cn^{-\frac{2k}{2k+2a+1}}$ provided $\tau \geq 4$. Therefore,

$$\widetilde{\Delta}_1 \leq C\Delta(n,\sigma,h_{opt}). \tag{2.43}$$

Now we find an upper bound for $\widetilde{\Delta}_2$. For $\widetilde{\Delta}_2$, $\hat{h} \geq h_{opt}$. Recall that, by definition of $\hat{h}$, if $\hat{h} = h \geq h_{opt}$, then

$$\|\hat{f}_{W,h} - \hat{f}_{W,h_{opt}}\|^2 \leq 16\, C^2(\tau,q)\, I(\sigma,h_{opt})\, n^{-1}\, \ln n.$$

Therefore,

$$\begin{aligned}\widetilde{\Delta}_2 &\leq \mathbb{E}\left[\|\hat{f}_{W,\hat{h}} - f_W\|^2 I(\hat{h} \geq h_{opt})\right] \\ &\leq \mathbb{E}\left[\|\hat{f}_{W,\hat{h}} - \hat{f}_{W,h_{opt}}\|^2\; I(\hat{h} \geq h_{opt})\right] + 2\mathbb{E}\|\hat{f}_{W,h_{opt}} - f_W\|^2 \\ &\leq 32C^2(\tau,q)n^{-1}\, I(\sigma,h_{opt}) \ln n + \Delta(n,\sigma,h_{opt}) \leq C\Delta(n,\sigma,h_{opt})\; \ln n,\end{aligned}$$

where $\Delta(n,\sigma,h)$ is defined in (2.17). Hence,

$$\widetilde{\Delta}_2 \leq C\Delta(n,\sigma,h_{opt})\; \ln n \tag{2.44}$$

Now we find an upper bound for $\widetilde{\Delta}_3$. Note that

$$\widetilde{\Delta}_3 \leq \sum_{j=j_{opt}+1}^{j_{\max}} \mathbb{E}\left[\|\hat{f}_{W,h} - f_W\|^2 I(\hat{h} = h = 2^{-j})\right]$$

If $\hat{h} = h = 2^{-j}$ for $j \geq j_{opt} + 1$, then $\hat{h} < h_{opt}$ and, by the definition of $\hat{h}$, there exists $\tilde{j}$ and $\tilde{h} = 2^{-\tilde{j}} < h_{opt}$, such that

$$\|\hat{f}_{W,h_{opt}} - \hat{f}_{W,\tilde{h}}\|^2 \geq 16\, C^2(\tau, q)\, I(\sigma, \tilde{h})\, n^{-1}\, \ln n. \tag{2.45}$$

Since for any $h \leq h_{opt}$,

$$\|f_{W,h} - f_W\|^2 \leq C_0\, n^{-1} I(\sigma, h)$$

where C_0 is an absolute constant, one has

$$\begin{aligned} \|\hat{f}_{W,h_{opt}} - \hat{f}_{W,\tilde{h}}\| &\leq \|\hat{f}_{W,h_{opt}} - f_{W,h_{opt}}\| + \|\hat{f}_{W,\tilde{h}} - f_{W,\tilde{h}}\| + \|f_{W,h_{opt}} - f_W\| + \|f_{W,\tilde{h}} - f_W\| \\ &\leq C_0\sqrt{n^{-1} I(\sigma, h_{opt})} + C_0\sqrt{n^{-1} I(\sigma, \tilde{h})} + \|\hat{f}_{W,h_{opt}} - f_{W,h_{opt}}\| + \|\hat{f}_{W,\tilde{h}} - f_{W,\tilde{h}}\|. \end{aligned}$$

Hence, by Lemma 2.7.1, if n is large enough,

$$\begin{aligned} &\mathbb{P}\left\{\|\hat{f}_{W,h_{opt}} - f_{W,\tilde{h}}\| \geq 4C(\tau, q)\sqrt{\tfrac{\ln n}{n} I(\sigma, \tilde{h})}\right\} \\ &\leq \mathbb{P}\left\{\|\hat{f}_{W,h_{opt}} - f_{W,h_{opt}}\| \geq 2C(\tau, q)\sqrt{\tfrac{\ln n}{n} I(\sigma, \tilde{h})} - C_0\sqrt{\tfrac{I(\sigma, h_{opt})}{n}}\right\} \\ &+\mathbb{P}\left\{\|\hat{f}_{W,\tilde{h}} - f_{W,\tilde{h}}\| \geq 2C(\tau, q)\sqrt{\tfrac{\ln n}{n} I(\sigma, \tilde{h})} - C_0\sqrt{\tfrac{I(\sigma, \tilde{h})}{n}}\right\} \\ &\leq \mathbb{P}\left\{\|\hat{f}_{W,h_{opt}} - f_{W,h_{opt}}\| \geq C(\tau, q)\sqrt{\tfrac{\ln n}{n} I(\sigma, h_{opt})}\right\} + \mathbb{P}\left\{\|\hat{f}_{W,\tilde{h}} - f_{W,\tilde{h}}\| \geq C(\tau, q)\sqrt{\tfrac{\ln n}{n} I(\sigma, \tilde{h})}\right\} \\ &\leq 2(2 + C_q)\, n^{-\tau}. \end{aligned}$$

Therefore

$$\widetilde{\Delta}_3 \leq \sum_{j=j_{opt}+1}^{j_{\max}} \sum_{\tilde{j}=j_{opt}+1}^{j_{\max}} \mathbb{E}\left[\|\hat{f}_{W,\hat{h}} - f_W\|^2 I(\hat{h} = 2^{-j}) I(\tilde{h} = 2^{-\tilde{j}})\right]$$

where $\tilde{h}$ is such that the inequality (2.45) holds. Let $\Omega_{\tilde{h}}$ be a set on which (2.45) is true. Then $\mathbb{P}(\Omega_{\tilde{h}}) \le 2(2 + C_q)\, n^{-\tau}$ and

$$\begin{aligned}
\widetilde{\Delta}_3 &\le \sum_{j=j_{opt}+1}^{j_{\max}} \sum_{\tilde{j}=j_{opt}+1}^{j_{\max}} \sqrt{\mathbb{E}\|\hat{f}_{W,\hat{h}} - f_W\|^4}\sqrt{\mathbb{P}(\Omega_{\tilde{h}})} I(\tilde{h} = 2^{-\tilde{j}})\, I(\hat{h} = 2^{-j}) \\
&\lesssim \sum_{j=j_{opt}+1}^{j_{\max}} \sum_{\tilde{j}=j_{opt}+1}^{j_{\max}} n^{1-\frac{\tau}{2}} \le C(\ln n)^2 n^{1-\frac{\tau}{2}} \le C\Delta(n, \sigma, h_{opt})
\end{aligned}$$

if $\tau/2 - 1 \ge 1$ which is true iff $\tau \ge 4$. Combination of the last inequality with (2.42), (2.43) and (2.44) complete the proof. □

2.7.5 Supplementary Lemmas and Their Proofs

Lemma 2.7.1. *Consider $Y_1, Y_2, \cdots, Y_n$ i.i.d such that $\mathbb{E}(|Y_1|^q) \le C_q$ with $q > 0$. Let $\tau \ge 1$ and $C(\tau, q)$ satisfies assumption* (2.26) *and $I(\sigma, h)$ be defined by* (2.16)*. Then there exists a set Ω such that for $w \in \Omega$ and all $h \ge 1/n$ simultaneously*

$$\|\hat{f}_{W,h}(x) - f_{W,h}(x)\|^2 \le C(\tau, q)^2\, I(\sigma, h)\, n^{-1} \ln n \tag{2.46}$$

and

$$\mathbb{P}(\Omega) \ge 1 - (2 + C_q)n^{-\tau} \tag{2.47}$$

Proof. Let $f_Y^*(w) = \mathbb{E}(\hat{f}_Y^*(w))$ where $\hat{f}_Y^*(w) = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} \exp(iY_k w) = \dfrac{1}{n}\sum_{k=1}^{n} \big[\cos(Y_k w) + i\sin(Y_k w)\big]$. First we show there exists a set Ω such that for $w \in \Omega$

$$\mathbb{P}\left(\sup_{|w| \le n} |\hat{f}_Y^*(w) - f_Y^*(w)| > C(\tau, q)\sqrt{\ln n/n}\right) \le (2 + C_q)n^{-\tau} \tag{2.48}$$

provided $C(\tau, q)$ satisfies assumption (2.26). Then it is sufficient to prove that

$$\mathbb{P}\left(\sup_{|w|\le n}\left|\frac{1}{n}\sum_{k=1}^{n}\left[\cos(Y_k w) - \mathbb{E}(\cos(Y_k w))\right]\right| > \frac{C(\tau,q)}{2}\sqrt{\frac{\ln n}{n}}\right) \le \frac{2+C_q}{n^{\tau}} \tag{2.49}$$

Let $\mathcal{B}$ be the set where the inequality (2.49) holds. For any $\gamma > 0$,

$$\mathbb{P}(\mathcal{B}) \le \mathbb{P}\left(\mathcal{B} \cap \{\max_{1\le k\le n}|Y_k| \le n^{\gamma}\}\right) + \mathbb{P}\left(\max_{1\le k\le n}|Y_k| > n^{\gamma}\right) \tag{2.50}$$

By Markov inequality,

$$\mathbb{P}\left(\max_{1\le k\le n}|Y_k| \ge n^{\gamma}\right) \le n^{-\gamma q}\,\mathbb{E}\left(\max_{1\le k\le n}|Y_k|^q\right) \le n^{-\gamma q}\sum_{k=1}^{n}\mathbb{E}|Y_k|^q \le n^{-\gamma q+1}\,\mathbb{E}|Y_1|^q. \tag{2.51}$$

Set $\gamma = (\tau+1)/q$, hence, $\gamma q - 1 = \tau$. Then

$$\mathbb{P}(\mathcal{B}) \le \mathbb{P}\left(\mathcal{B} \cap \{\max_{1\le k\le n}|Y_k| \le n^{\gamma}\}\right) + n^{-\tau}\,\mathbb{E}|Y_1|^q. \tag{2.52}$$

Partition the interval $[-n, n]$ into M sub-intervals by points $w_j, j = 0, 1, 2, 3, \cdots, M$, such that $w_0 = -n$, $w_j - w_{j-1} = n^{-(\gamma+1)}$, so that $M = 2n^{\gamma+2}$. Consider a random functions $Z_k(w) = \left[\cos(Y_k w) - \mathbb{E}(\cos(Y_k w))\right]\mathbb{I}(|Y_k| \le n^{\gamma})$. Since $|Y_k| \le n^{\gamma}$ and $|\partial(\cos(Yw))/\partial w| \le |Y| \le n^{\gamma}$, obtain

$$|Z_k(w) - Z_k(w')| \le 2n^{\gamma}|w - w'|.$$

Therefore $Z_k(w)$ satisfies the Lipschitz condition and, for any $w \in [-n, n]$, there exists w_j such that

$$\left|\frac{1}{n}\sum_{k=1}^{n}Z_k(w)\right| \le \left|\frac{1}{n}\sum_{k=1}^{n}Z_k(w_j)\right| + 2n^{\gamma}.\frac{1}{n^{\gamma+1}}$$

Hence,

$$\begin{aligned}
\mathbb{P}\left(\mathcal{B} \cap \{\max_{1\le k\le n}|Y_k| \le n^\gamma\}\right) &\le \mathbb{P}\left(\max_{1\le j\le M}\left|\frac{1}{n}\sum_{k=1}^{n} Z_k(w_j)\right| + \frac{2}{n} > \frac{C(\tau,q)}{2}\sqrt{\frac{\ln n}{n}}\right) \\
&\le \mathbb{P}\left(\max_{1\le j\le M}\left|\frac{1}{n}\sum_{k=1}^{n} Z_k(w_j)\right| > \frac{C(\tau,q)}{4}\sqrt{\frac{\ln n}{n}}\right) \\
&\le \sum_{j=1}^{M}\mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n} Z_k(w_j)\right| > \frac{C(\tau,q)}{4}\sqrt{\frac{\ln n}{n}}\right)
\end{aligned}$$

provided

$$\frac{C(\tau,q)}{4}\sqrt{\frac{\ln n}{n}} \ge \frac{2}{n},$$

which is guaranteed by condition (2.26).

Using Hoeffding inequality with $\xi_k = Z_k(w_j)$ where $|\xi_k| \le 2$ and $t = \frac{C(\tau,q)}{4}\sqrt{\frac{\ln n}{n}}$, obtain that

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n} Z_k(w_j)\right| > \frac{C(\tau,q)}{4}\sqrt{\frac{\ln n}{n}}\right) \le 2\exp\left(-\frac{(C(\tau,q))^2 \ln n}{128}\right)$$

and

$$\mathbb{P}\left(\mathcal{B} \cap \{\max_{1\le k\le n}|Y_k| \le n^\gamma\}\right) \le 2n^{-\tau} \tag{2.53}$$

is guaranteed by condition (2.26). Validity of (2.48) follows from the inequality (2.52) and (2.53).

In order to prove (2.46), note that $1/h \le n$ and

$$\begin{aligned}
\|\hat{f}_{W,h} - f_{W,h}\|^2 = & \ \tfrac{1}{2\pi} \|\hat{f}^*_{W,h} - f^*_{W,h}\|^2 \\
& = \tfrac{1}{2\pi} \int_{-1/h}^{1/h} \frac{|g^*(\sigma s)|^2}{|f^*_\xi(s)|^2} |\hat{f}^*_Y(s) - f^*_Y(s)|^2 ds \\
& \le \sup_{|s| \le n} |\hat{f}^*_Y(s) - f^*_Y(s)|^2 \, I(\sigma, h)
\end{aligned}$$

which completes the proof.

□

Lemma 2.7.2. *Let* $h_{\min} = \max\{\sigma, n^{-\frac{1}{2a+1}}\}$. *Then, for any* $h \in [h_{\min}, 1/2]$ *, one has*

$$\mathbb{E}\|\hat{f}_{W,h} - f_W\|^4 \le \begin{cases} C\sigma^{-(4a+3)} n^{-1}, & \textit{cases I, IV} \\ C\sigma^{-(4a+3)} n^{-1} \ln n, & \textit{case II} \\ Cn^2, & \textit{case III} \end{cases}$$

In particular, if $\sigma \ge n^{-\frac{1}{2a+1}}$, *then* $\mathbb{E}\|\hat{f}_{W,h} - f_W\|^4 \le Cn^2$.

Proof. Note that

$$\mathbb{E}\|\hat{f}_{W,h} - f_W\|^4 = \mathbb{E}\|\hat{f}_{W,h} - \mathbb{E}\hat{f}_{W,h} + \mathbb{E}\hat{f}_{W,h} - f_W\|^4 \le 8\mathbb{E}\|\hat{f}_{W,h} - \mathbb{E}\hat{f}_{W,h}\|^4 + 8\|\mathbb{E}\hat{f}_{W,h} - f_W\|^4. \tag{2.54}$$

Then,

$$\|\mathbb{E}\hat{f}_{W,h} - f_W\|^4 = [R_1(\hat{f}_{W,h}, f_W)]^2 \le \Delta_1^2 \le 1$$

where Δ_1 is defined in (2.19). To find an upper bound for the first term, note that for any x

$$\begin{aligned}
|\hat{f}_{W,h}(x)| &\leq \tfrac{1}{2\pi} \int_{-1/h}^{1/h} \frac{|g^*(\sigma s)|}{|f_\xi^*(s)|}\, ds \\
&\leq \tfrac{1}{2\pi} \int_{-1/h}^{1/h} \frac{C_g(\sigma^2 s^2+1)^{-\frac{\vartheta}{2}} \exp(-\gamma|\sigma s|^\beta)}{c_\xi (s^2+1)^{-\frac{a}{2}}}\, ds \\
&\leq C \min\left(h^{-(a+1)}, \sigma^{-(a+1)}\right) \tilde{\Delta}_{h,\sigma}
\end{aligned}$$

where

$$\tilde{\Delta}_{h,\sigma} = \begin{cases} 1 & \text{in case I and IV} \\ \max\left\{\ln\left(\frac{\sigma}{h}\right), 1\right\} & \text{in case II} \\ \max\left\{1, \left(\frac{\sigma}{h}\right)^{a-\vartheta+1}\right\} & \text{in case III} \end{cases}$$

The same upper bound holds for $f_{W,h} = \mathbb{E}\hat{f}_{W,h}$. Hence,

$$\|\hat{f}_{W,h} - \mathbb{E}\hat{f}_{W,h}\|_\infty^2 \leq C \min\left(h^{-2(a+1)}, \sigma^{-2(a+1)}\right) \tilde{\Delta}_{h,\sigma}^2. \tag{2.55}$$

Therefore,

$$\begin{aligned}
\mathbb{E}\|\hat{f}_{W,h} - f_{W,h}\|^4 &\leq \mathbb{E}\left[\|\hat{f}_{W,h} - f_{W,h}\|^2\right] \|\hat{f}_{W,h} - f_{W,h}\|_\infty^2 \\
&\leq C\, n^{-1} \min\left(h^{-(4a+3)}, \sigma^{-(4a+3)}\right) \tilde{\Delta}_{h,\sigma}^2 \Delta_{h,\sigma}^2
\end{aligned}$$

where, according to Lemma 2.4.2, $\Delta_{h,\sigma}$ is of the form (2.35). Observe that an upper bound for the first term in (2.54) is larger than the second term and that

$$\mathbb{E}\|\hat{f}_{W,h} - f_W\|^4 \leq \begin{cases} C\sigma^{-(4a+3)} n^{-1}, & \text{in cases I, IV} \\ C\sigma^{-(4a+3)} n^{-1} \ln\left(\frac{1}{h_{\min}}\right), & \text{in case II} \\ h^{-(4a+3)} n^{-1}, & \text{in case III} \end{cases}$$

Since $h_{\min} \geq n^{-\frac{1}{2a+1}}$, we finally obtain (2.54).

Now, let $\sigma \geq n^{-\frac{1}{2a+1}}$, then $\sigma^{-(4a+3)} n^{-1} \leq n^{-1} n^{\frac{4a+3}{2a+1}} \leq n^{\frac{2a+2}{2a+1}} \leq n^2$, which completes the proof.

□

Lemma 2.7.3. *Let* $\sigma \leq n^{-\frac{1}{2a+2k+1}}$ *and* $\Omega_{\sigma,n}$ *be defined by formula* (2.27). *Then in the cases I and IV, if* n *is large enough,*

$$\mathbb{P}(\Omega_{\sigma,n}) \leq (2 + C_q)\, n^{-\tau} \tag{2.56}$$

Proof. Note that

$$\|\hat{f}_{W,\sigma} - \hat{f}_{W,\frac{1}{n}}\| \leq \ \|\hat{f}_{W,\sigma} - f_{W,\sigma}\| + \|\hat{f}_{W,\frac{1}{n}} - f_{W,\frac{1}{n}}\| + \|f_{W,\sigma} - f_W\| + \|f_{W,\frac{1}{n}} - f_W\| \tag{2.57}$$

Then by Lemma 2.4.1, for some absolute constant $\tilde{C}$,

$$\|f_{W,\sigma} - f_W\| \leq \tilde{C}\sigma^{-k};\ \|f_{W,\frac{1}{n}} - f_W\| \leq \tilde{C}n^{-k} < \tilde{C}\sigma^k$$

Also, by Corollary 2.7.1, for $w \in \Omega$

$$\|\hat{f}_{W,\sigma} - f_{W,\sigma}\| \leq C(\tau, q)\sqrt{\frac{I(\sigma,\sigma)\ln n}{n}} \leq C(\tau, q)\sqrt{\frac{I(\sigma, 1/n)\ln n}{n}}$$

and

$$\|\hat{f}_{W,\frac{1}{n}} - f_{W,\frac{1}{n}}\| \leq C(\tau, q)\sqrt{\frac{I(\sigma, 1/n)\ln n}{n}}$$

Hence, it follows from (2.57), that for $w \in \Omega$

$$\|\hat{f}_{W,\sigma} - \hat{f}_{W,\frac{1}{n}}\| \leq 2C(\tau, q)\sqrt{\frac{I(\sigma, 1/n)\ln n}{n}} + 2\tilde{C}\sigma^k.$$

Therefore, for $w \in \Omega$,

$$\|\hat{f}_{W,\sigma} - \hat{f}_{W,\frac{1}{n}}\| \geq 4C(\tau, q)\sqrt{\frac{I(\sigma, 1/n)\ln n}{n}}$$

cannot be true, unless

$$C(\tau,q)\sqrt{\frac{I(\sigma,1/n)\ln n}{n}} < \tilde{C}\sigma^k. \tag{2.58}$$

By Lemma 2.4.2, in the cases I and IV, one has $I(\sigma,1/n) \le \tilde{\tilde{C}}\sigma^{-(2a+1)}$. So, inequality (2.58) holds only if $C(\tau,q)\ (\tilde{\tilde{C}})^2\sigma^{-(a+\frac{1}{2})}\sqrt{\ln n/n} < \tilde{C}\sigma^k$, which is equivalent to $\sigma > \bar{C}\left(n^{-1}\ln n\right)^{\frac{1}{2k+2a+1}}$ where $\bar{C} = \tilde{\tilde{C}}C(\tau,q)/\tilde{C}$. Therefore, if $w \in \Omega$ and $\sigma \le n^{-\frac{1}{2k+2a+1}}$, where n is such that $\bar{C}\,(\ln n)^{\frac{1}{2k+2a+1}} \ge 1$, then (2.58) is not true. Hence, $w \notin \Omega_{\sigma,n}$, so that $\Omega_{\sigma,n} \subset \Omega^c$ and (2.56) holds.

□

Lemma 2.7.4. *Consider an integral of the form*

$$I(\lambda) = \int_{m_1}^{m_2} P_\lambda(z)\exp(Q_\lambda(z))dz \tag{2.59}$$

where $0 \le m_1 < m_2 < \infty$ and $P_\lambda(z)$ and $Q_\lambda(z)$ are real valued differentiable functions of z and $\lambda \to \infty$ is a large parameter. Let

$$z_0 \equiv z_{0,\lambda} = \underset{z\in[m_1,m_2]}{\mathit{argmax}}\, Q_\lambda(z)$$

be an unique global maximum of $Q_\lambda(z)$ on the interval $[m_1,m_2]$. Assume that the following conditions hold:

- *A function P is a positive slowly varying function, i.e., for any $t > 0$ one has $\lim_{x\to\infty} P(tx)/P(x) = 1$.*
- *$Q_\lambda(z_0) - Q_\lambda(z)$ increases monotonically for $\lambda \ge \lambda_0$ as $\lambda \to \infty$.*

- *If* $Q'_\lambda(z_0) = 0$, *then for every* $\lambda \geq \lambda_0$

$$\lim_{x\to 0} \frac{Q_\lambda(z_0+x) - Q_\lambda(z_0)}{x^2} = \frac{Q''_\lambda(z_0)}{2} < 0 \tag{2.60}$$

- *If* $Q'_\lambda(z_0) \neq 0$, *then for every* $\lambda \geq \lambda_0$

$$\lim_{x\to 0} \frac{Q_\lambda(z_0+x) - Q_\lambda(z_0)}{x} = Q'_\lambda(z_0) \neq 0 \tag{2.61}$$

Then, as $\lambda \to \infty$,

$$I(\lambda) \asymp \begin{cases} \dfrac{\exp\{Q_\lambda(z_0)\}\, P_\lambda(z_0)}{\sqrt{|Q''_\lambda(z_0)|}}, & \textit{if} \quad \text{(2.60)}\ \textit{holds}, \\ \dfrac{\exp\{Q_\lambda(z_0)\}\, P_\lambda(z_0)}{Q'_\lambda(z_0)}, & \textit{if} \quad \text{(2.61)}\ \textit{holds} \end{cases} \tag{2.62}$$

Proof. Comparing (2.59) with the integral

$$I(\lambda) = \int G(z) \exp(-F(z)) dz \tag{2.63}$$

obtain $F(z) = -Q_\lambda(z), G(z) = P_\lambda(z)$. Then, following the calculations in Dingle (1973) with $F(z_0) = -Q_\lambda(z_0), F_1(z_0) = -Q'_\lambda(z_0)$, from the formulas (3) and (4), page 111, obtain

$$I(\lambda) = [-Q'_\lambda(z_0)]^{-1} \exp\{Q_\lambda(z_0)\} \sum_0^\infty L_r$$

where L_r is given by

$$L_r = -Q'_\lambda(z_0) \left(\frac{d}{Q'_\lambda(z)\, dz} \right)^r \frac{P_\lambda(z)}{Q'_\lambda(z)} \Bigg|_{z=z_0}$$

Hence, taking the term with $r = 0$, obtain, when (2.60) holds:

$$I(\lambda) \approx \frac{\exp\{Q_\lambda(z_0)\}\, P_\lambda(z_0)}{Q'_\lambda(z_0)}.$$

Now, consider the case when (2.61) holds.

Then following the calculations in Dingle (1973), page 118, obtain

$$I(\lambda) = \exp\{Q_\lambda(z_0)\} \int \exp\{-f^2\} P_\lambda(z) dz$$

where

$$f = \sqrt{Q_\lambda(z_0) - Q_\lambda(z)} \sim \sqrt{F_2/2}\, z \quad \text{as} \quad z \to z_0$$

with $F_2(z_0) = -Q''_\lambda(z_0)$. Therefore, from formulas (16) and (17), page 119, obtain

$$I(\lambda) = \left[\frac{\pi}{-2Q''_\lambda(z_0)}\right]^{\frac{1}{2}} \exp\{Q_\lambda(z_0)\} \sum_0^\infty L_r$$

where L_r is given by

$$L_r = Q'_\lambda(z_0) \left(\frac{d}{2\, f'\, dz}\right)^r \frac{P_\lambda(z)}{f'}\Bigg|_{z=z_0}$$

Hence, taking the term with $r = 0$, obtain, when (2.61) holds:

$$I(\lambda) \approx \frac{\sqrt{\pi}\exp\{Q_\lambda(z_0)\}\, P_\lambda(z_0)}{\sqrt{-2Q''_\lambda(z_0)}}$$

which is equivalent to second expression of (2.62).

□

Lemma 2.7.5. *Let n be large and $z \in \mathbb{R}$ be a fixed quantity. Then, as $n \to \infty$, the solution of the* *equation*

$$e^m m^z = n \tag{2.64}$$

is given by

$$m = (\ln n - z \ln\ln n)(1 + o(1)), \quad n \to \infty. \tag{2.65}$$

Proof. Since $e^m m^z = n$, then $m + z \ln m = \ln n$ and $m = \ln n - z \ln m$. Plugging this m back into (2.64), obtain $e^{\ln n - z \ln m}(\ln n - z \ln m)^z = n$. Since for large values of n, one has $(\ln n - z \ln m)^z \approx (\ln n)^z$, the previous equation becomes $(\ln n)^z n e^{-z \ln m} \approx n$, so that $z \ln\ln n \approx z \ln m$ which yields (2.65).

□

CHAPTER 3: ESTIMATION AND CLUSTERING IN THE POPULARITY ADJUSTED BLOCK MODEL

3.1 Statistical Network Models: Background

We can see that networks are available everywhere in science. They have become a center of attention for discussion in everyday life. With the evolution of digital technology, the network data are more readily available than before. Since many fields involve the study of networks in some form, the formal statistical models for the analysis of network data have emerged as a major topic of interest in diverse areas of study. Networks have been used in predicting community evolution, recommendation systems (for example: Netflix and Youtube in recommending movies), targeted marketing (for example: Amazon suggesting items), personal influence (for example: politics, link prediction), detection of disease (for example: cancer or tumor types), criminology (for example: to identify criminal group, fraud, etc) to name a few. The existing set of statistical network models may be classified into various categories, but we only consider the one that is related to the static network. Static network models concentrate on explaining the observed set of links based on a single snapshot of the network (Goldenberg et al. (2010)). We start with the definition of the random network.

3.1.1 Random Network

As it was said, a network $G = (V, E)$ is a structure made of nodes (vertices) denoted by V and edges (also called links) denoted by E, that connect nodes in various relationships. Networks are more commonly represented as graphs. They are also represented in terms of matrices known as adjacency matrices. Networks can be weighted, signed, undirected, and directed. The edges in the weighted network are associated with numerical values. Edges in signed network are associated with positive and negative relationships. Directed networks have direction associated with edges whereas direction does not matter in undirected network. The degree of the nodes represents the number of links it has to other nodes.

A random network consists of n nodes where each pair of nodes is connected with the probability p. The goal of the random network is to build a model that reproduces the properties of the real networks (Barabási et al. (2016)). To construct a random network (also called a random graph), we start with the n isolated nodes, then select a pair of nodes and generate a random number between 0 and 1. If a generated number is at least p, connect the selected node pair with a link, otherwise leave them disconnected. We continue this process for each of the $\binom{n}{2}$ pairs of nodes. A random network is called the Erdős-Rényi network, in honor of mathematicians Pál Erdős and Alfréd Rényi, who played an important role in understanding their properties.

Table 3.1: Examples of some real networks

Network	Node	Link	Network Type
I) Citation Network	Paper	Citations	Directed
II) Email	Email Addresses	Emails	Directed
III) Internet	Routers	Connections	Undirected
IV) Social Network	Users	Interactions	Undirected
IV)Coauthorship Network	Research Scientist	Coauthor a paper	Undirected

In this dissertation, we work with the undirected network. For this type of network, the ij^{th} element of the adjacency matrix is defined as 1 if there is an edge between the i^{th} node and the j^{th} node, and zero otherwise. The adjacency matrix is symmetric in this case.

3.2 Community Structure and Blockmodels

Community detection aims at finding clusters as subgraphs within a given network. A community is a cluster where many edges link nodes of the same group and few edges link nodes of different group. We follow the general approach to community detection by considering a network as a static view in which all the nodes and links in the network are kept unchanged throughout the study. We assume that the network under consideration have underlying blocks. The assumption in the block model is that the nodes inside the blocks have higher connectivity in comparison to the nodes between the blocks. There are various random graph model on the community detection problems such as Stochastic Block Model, Degree-Corrected Block Model, and Popularity Adjusted Stochastic Block Model and its variants. We study these models and developed the novel estimation procedure for the Popularity Adjusted Stochastic Block Model and implemented the Sparse Subspace Clustering method for the community detection in the model.

3.2.1 Stochastic Block Model (SBM)

Consider an undirected network with n nodes and no self loops and multiple edges. Let A be its adjacency matrix. Then $A_{ij} = \text{Ber}(P_{ij})$, where P is a symmetric probability matrix whose diagonal entries are zeroes and whose off-diagonal entries are between 0 and 1. A classical random graph model for networks with community structure is the SBM that was introduced by Lorrain and White (1971), Holland et al. (1983) and E. Fienberg et al. (1985). Under this model, all nodes belonging to a community are considered to be stochastically equivalent, in the sense that they have the same probability of forming a link with another node in the network. Various methods of community detection have been studied under the SBM, examples include spectral clustering Rohe et al. (2011), variational methods Celisse et al. (2012), and pseudo-likelihood methods Amini et al. (2013).

Under the K-block SBM, each node belongs to one of K distinct blocks or communities. Let c denote the true community assignment vector with $c_i = a$ if the i^{th} node belongs to the a^{th} community. Then for $i < j, P_{ij} = B_{c_i c_j}$ where $B_{k,l}$ is the probability of connection between communities k and l. and so B is the K-by-K matrix of community link probabilities. Edges are conditionally independent given c and B.

Under the SBM, two nodes belonging to the same community display community structure by behaving identically, in a stochastic sense. In particular, any two nodes from the same community have the same degree distribution and the same expected degree. Since the real-life networks usually contain a very small number of high-degree nodes while the rest of the nodes have very few connections (low degree), the SBM model fails to explain the structure of many networks that occur in practice.

3.2.2 Degree Corrected Block Model (DCBM)

DCBM introduced by Karrer and Newman (2011) adds node-specific degree parameters such that for $i < j, P_{ij} = \theta_i B_{c_i c_j} \theta_j$, where θ_i and θ_j are the degree parameters for the respective nodes, and B is the K-by-K matrix of baseline interaction between communities. Edges are conditionally independent given c, θ, and B. Identifiability of the parameters is ensured by a constraint of the form $\sum_{i \in \mathcal{N}_a} \theta_i = 1, \forall a = 1, ..., K$, where $\mathcal{N}_a$ is the set of nodes belonging to community a. DCBM enforces node popularity to be uniformly proportional to the node degree. DCBM correctly detects the communities, and accurately fits the total degree, by enforcing the node-specific degree parameters. However the model fitting for node popularity is quite inaccurate.

3.2.3 The Popularity Adjusted Block Model (PABM)

A network feature that is closely associated with community structure is the popularity of nodes across communities, defined as the number of edges between a specific node and a specific community. Node popularity is inseparably associated with community structure. The observed popularity of the i^{th} node in the r^{th} community is given by $M_{ir} = \sum_{i \in \mathcal{N}_r} A_{ij}$ and the expectation of this quantity is defined as $\mu_{ir} = \mathbb{E}[M_{ir}] = \sum_{i \in \mathcal{N}_r} P_{ij}$. We called μ_{ir} is the popularity of the i^{th} node in the r^{th} community. In practice, observed popularities of the n nodes in the K communities vary considerably across nodes along with communities. SBM and the DCBM both put unrealistic restrictions on node popularities.

DCBM fails to model node popularities in a flexible and realistic way. To fulfill the need of the model that allows flexible and realistic modeling of node popularity, Sengupta and Chen (2018) introduced a new random graph model, called PABM for modeling node popularity in networks with community structure. They developed methodology for community detection and parameter estimation under the PABM, and demonstrated the improvement achieved through this new methodology. In PABM, for $i < j$,

$$P_{ij} = V_{ic_j} V_{jc_i} \tag{3.1}$$

where $V_{ir}, 1 \leq i \leq n, 1 \leq r \leq K$, are the popularity scaling parameters and $0 \leq P_{ij} \leq 1$ for all $i < j$. Thus, P_{ij} depends on the popularity of node i in the community to which j belongs, and the popularity of node j in the community to which i belongs. Similar to the identifiability issue with the DCBM as discussed in Karrer and Newman (2011), the PABM also has a scaling identifiability issue. To resolve that issue, Sengupta and Chen (2018) impose the identifiability constraint $V_{rs} = V_{sr}$ where $V_{rs} = \sum\limits_{j \in \mathcal{N}_r} V_{js}$.

Furthermore, Set $V_{ir} = \theta_i \sqrt{B_{c_i r}}$ so that

$$V_{ic_j} = \theta_i \sqrt{B_{c_i c_j}}$$

and

$$V_{jc_i} = \theta_j \sqrt{B_{c_i c_j}}$$

Then

$$P_{ij} = V_{ic_j} V_{jc_i} = \theta_i \sqrt{B_{c_i c_j}}\, \theta_j \sqrt{B_{c_i c_j}} = \theta_i B_{c_i c_j} \theta_j$$

which shows that DCBM is a special case of PABM. Also if you set $\theta_i = 1$ in the definition of V_{ir} above, so that

$$V_{ic_j} = \sqrt{B_{c_i c_j}}$$

and

$$V_{jc_i} = \sqrt{B_{c_i c_j}}$$

Then

$$P_{ij} = V_{ic_j} V_{jc_i} = \sqrt{B_{c_i c_j}}\, \sqrt{B_{c_i c_j}} = B_{c_i c_j}$$

which shows that SBM is a special case of PABM.

3.3 PABM: the Structure of the Probability Matrix

The ratio of popularities of the nodes $(i, j) \in \mathcal{N}_k$ in the same community k is equal to one for the SBM, is independent of community k (a function of i and j only) in DCBM but can vary between nodes and communities for the PABM, thus, allowing a more flexible modeling of connection probabilities. Heuristically, the degree of node is a network-level feature, DCBM can model this feature well by allowing each node to have its own degree parameter. Popularity of the node is the community level feature since the same node can be popular in one community and unpopular in the other community. DCBM cannot model this feature accurately because it governs the relative behavior of a node in all communities by a single degree parameter which forces a high degree node to be relatively popular uniformly across the network and low degree node to be uniformly unpopular. PABM fixes this issue by assigning parameters for every node- community combination.

The flexibility of PABM, however, is not limited to modeling the popularity parameters of the nodes. In order to better understand the model, consider a rearranged version $P(Z, K)$ of matrix P where its first n_1 rows correspond to nodes from class 1, the next n_2 rows correspond to nodes from class 2 and the last n_K rows correspond to nodes from class K. Denote the (k, l)-th block of matrix $P(Z, K)$ by $P^{(k,l)}(Z, K)$. Since sub-matrix $P^{(k,l)}(Z, K) \in [0, 1]^{n_k \times n_l}$ corresponds to pairs of nodes in communities (k, l) respectively, one obtains from (3.1) that $P^{(k,l)}_{i,j} = V_{i_k,l} V_{j_l,k}$ where i_k is the i-th element in $\mathcal{N}_k$ and j_l is the j-th element in $\mathcal{N}_l$. Thus, matrices $P^{(k,l)}(Z, K)$ are rank-one matrices with the unique singular vectors generating them. Indeed, consider vectors $\Lambda^{(k,l)}$ with elements $\Lambda^{(k,l)}_i = V_{i_k,l}$, where $i = 1, \dots, n_k$ and $i_k \in \mathcal{N}_k$. Then, equation (3.1) implies that

$$P^{(k,l)}(Z, K) = \Lambda^{(k,l)} \, [\Lambda^{(l,k)}]^T. \tag{3.2}$$

Moreover, it follows from (3.1) and (3.2) that $P^{(k,l)}(Z, K) = [P^{(l,k)}(Z, K)]^T$ and that each pair of blocks (k, l) involves a unique combination of vectors $\Lambda^{(l,k)}$:

$$P(Z, K) = \begin{bmatrix} \Lambda^{(1,1)}(\Lambda^{(1,1)})^T & \Lambda^{(1,2)}(\Lambda^{(2,1)})^T & \dots & \Lambda^{(1,K)}(\Lambda^{(K,1)})^T \\ \Lambda^{(2,1)}(\Lambda^{(1,2)})^T & \Lambda^{(2,2)}(\Lambda^{(2,2)})^T & \dots & \Lambda^{(2,K)}(\Lambda^{(K,2)})^T \\ \vdots & \vdots & \dots & \vdots \\ \Lambda^{(K,1)}(\Lambda^{(1,K)})^T & \Lambda^{(K,2)}(\Lambda^{(2,K)})^T & \dots & \Lambda^{(K,K)}(\Lambda^{(K,K)})^T \end{bmatrix} \tag{3.3}$$

where

$$\Lambda = \begin{bmatrix} \Lambda^{(1,1)} & \Lambda^{(1,2)} & \dots & \Lambda^{(1,K)} \\ \Lambda^{(2,1)} & \Lambda^{(2,2)} & \dots & \Lambda^{(2,K)} \\ \vdots & \vdots & \dots & \vdots \\ \Lambda^{(K,1)} & \Lambda^{(K,2)} & \dots & \Lambda^{(K,K)} \end{bmatrix} \tag{3.4}$$

The latter implies that matrix $P(Z, K)$ is formed by arbitrary rank one blocks and hence $rank(P(Z, K)) = rank(P)$ can take any value between K and K^2. In comparison, all other block models restrict the rank of P to be exactly K. This is true not only for the SBM and DCBM discussed above but also for their generalizations such as the Mixed Membership models (see, e.g., Airoldi et al. (2008) and Cheng et al. (2017)) and the Degree Corrected Mixed Membership (DCMM) (see, e.g., Jin et al. (2017)). Hence, the PABM allows for much more flexible spectral structure than any other block model above.

This flexibility makes the PABM an attractive choice for modeling networks that appear in biological sciences. Indeed, while social networks exhibit assortative behavior due to the human tendency of forming strong associations, the biological networks tend to be more diverse. For this reason, PABM tends to be a useful tool for modeling such networks.

However, while the PABM model is extremely valuable, the statistical inference in Sengupta and Chen (2018) has been incomplete. In particular, the authors considered only the case of a small finite number of communities K; they provided only asymptotic consistency results as $n \to \infty$ without any error bounds when n is finite; their clustering procedure was tailored to the case of a small K, therefore, all simulations and real data examples in Sengupta and Chen (2018) only tackled the case of $K = 2$.

The purpose of the present work is to address some of those deficiencies and to advance the theory of the PABM. Specifically, we make the following contributions:

1. In contrast to Sengupta and Chen (2018), we consider the PABM with an arbitrary number of communities which possibly grows with a number of nodes in the network and is not assumed to be known.
2. We argue that the main appeal of the PABM is the flexibility of the spectral properties of the graph and replace the estimators in Sengupta and Chen (2018) that are based on averaging over the communities by more accurate counterparts based on low rank matrix approximations.
3. While Sengupta and Chen (2018) only proved convergence of the estimation and clustering errors to zero as the number of nodes grows, we derive non-asymptotic upper bounds for those errors when the number of communities is arbitrary. In particular, we produce an upper bound for the estimation error of the matrix of the connection probabilities and provide a condition that guarantees that the proportion of misclassified nodes is bounded above by a specified quantity. All results in the PABM are non-asymptotic and are valid for any combination of parameters.

In the next chapter we discuss estimation and clustering in PABM as a solution of a penalized optimization procedure. We start with notations used throughout the chapter, formulate estimation and clustering as solutions of an optimization procedure, and derive upper bounds for estimation errors as well as find sufficient conditions for the proportion of misclustered nodes to be bounded above by a pre-specified quantity ρ_n with a high probability.

3.4 Notation

For any two positive sequences $\{a_n\}$ and $\{b_n\}$, $a_n \asymp b_n$ means that there exists a constant $C > 0$ independent of n such that $C^{-1} a_n \leq b_n \leq C a_n$ for any n. For any set Ω, denote cardinality of Ω by $|\Omega|$. For any numbers a and b, $a \wedge b = \min(a, b)$. For any vector $t \in \mathbb{R}^p$, denote its ℓ_2, ℓ_1, ℓ_0 and ℓ_∞ norms by, respectively, $\|t\|$, $\|t\|_1$, $\|t\|_0$ and $\|t\|_\infty$. Denote by 1_m the m-dimensional column vector with all components equal to one.

For any matrix A, denote its spectral and Frobenius norms by, respectively, $\|A\|_{op}$ and $\|A\|_F$. Let $\text{vec}(A)$ be the vector obtained from matrix A by sequentially stacking its columns.

Denote by $\Pi_J(X)$, the projection of a matrix $X : n \times m$ onto the set of matrices with non zero elements in the set $J = J_1 \times J_2 = \{(i, j) : i \in J_1,\ j \in J_2\}$. Denote by $\Pi_{(1)}(X)$ the best rank one approximation of matrix X and by $\Pi_{u,v}(X)$ the rank one projection of X onto pair of unit vectors u, v given by

$$\Pi_{u,v}(X) = (uu^T)X(vv^T). \tag{3.5}$$

Then, $\Pi_{(1)}(X) = \Pi_{u,v}(X)$ provided (u, v) is a pair of singular vectors of X corresponding to the largest singular value.

Denote by $\mathcal{M}_{n,K}$ a collection of clustering matrices $Z \in \{0, 1\}^{n \times K}$ such that $Z_{i,k} = 1$ iff $i \in \mathcal{N}_k$, $i = 1, \ldots, n$, and $Z^T Z = \text{diag}(n_1, \ldots, n_K)$ where $n_k = |\mathcal{N}_k|$ is the size of community k, where $k = 1, \ldots, K$. Denote by $\mathscr{P}_{Z,K} \in \{0, 1\}^{n \times n}$ the permutation matrix corresponding to $Z \in \mathcal{M}_{n,K}$ that rearranges any matrix $B \in R^{n,n}$, so that its first n_1 rows correspond to nodes from class 1, the next n_2 rows correspond to nodes from class 2 and the last n_K rows correspond to nodes from class K. Recall that $\mathscr{P}_{Z,K}$ is an orthogonal matrix with $\mathscr{P}_{Z,K}^{-1} = \mathscr{P}_{Z,K}^T$. For any $\mathscr{P}_{Z,K}$ and any matrix $B \in \mathbb{R}^{n \times n}$ denote the permuted matrix and its blocks by, respectively, $B(Z,K)$ and $B^{(k,l)}(Z,K)$, where $B^{(k,l)}(Z,K) \in \mathbb{R}^{n_k \times n_l}$, $k,l = 1,\dots,K$, and

$$B(Z,K) = \mathscr{P}_{Z,K}^T B \mathscr{P}_{Z,K}, \qquad B = \mathscr{P}_{Z,K} B(Z,K) \mathscr{P}_{Z,K}^T. \tag{3.6}$$

Also, in the present and the next chapter of the dissertation, we use the star symbol to identify the true quantities. In particular, we denote the true matrix of connection probabilities by P_*, the true number of classes by K_* and the true clustering matrix that partitions n nodes into K_* communities by Z_*.

3.5 Optimization Procedure for Estimation and Clustering

In this section we consider estimation of the true probability matrix P_*. Consider block $P_*^{(k,l)}(Z_*,K_*)$ of the rearranged version $P_*(Z_*,K_*)$ of P_*. Let $\Lambda \equiv \Lambda(Z_*,K_*) \in [0,1]^{n \times K_*}$ be a block matrix with each column l partitioned into K_* blocks $\Lambda^{(k,l)} \equiv \Lambda^{(k,l)}(Z_*,K_*) \in [0,1]^{n_k}$. Then, due to (3.2), $P_*^{(k,l)}(Z_*,K_*)$ are rank-one matrices such that $P_*^{(k,l)}(Z_*,K_*) = [P_*^{(l,k)}(Z_*,K_*)]^T$ and that each pair of blocks (k,l) involves a unique combination of vectors $\Lambda^{(k,l)}$. The structures of matrices $P_*(Z_*,K_*)$, Λ and P_* are illustrated in Figure 3.1.

Observe that although matrices $P_*^{(k,l)}(Z_*,K_*)$ in (3.2) are well defined, vectors $\Lambda^{(k,l)}$ and $\Lambda^{(l,k)}$ can be determined only up to a multiplicative constant. In particular, under the constraint

$$1_{n_k}^T \Lambda^{(k,l)} = 1_{n_l}^T \Lambda^{(l,k)}, \tag{3.7}$$

Sengupta and Chen (2018) obtained explicit expressions for vectors $\Lambda^{(k,l)}$ and $\Lambda^{(l,k)}$ in (3.2).

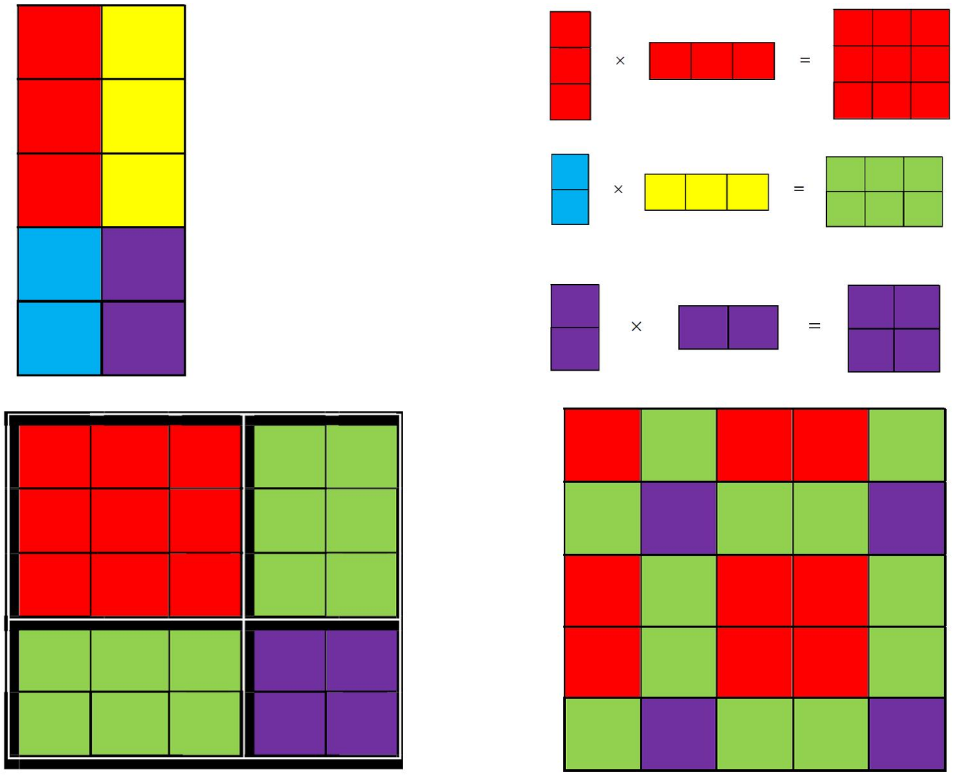

Figure 3.1: Matrices Λ, $P(Z,K)$ and P in the case of $n = 5$ and $K = 2$. Matrix Λ (top left): $\Lambda^{(1,1)}$ (red), $\Lambda^{(2,1)}$ (blue), $\Lambda^{(1,2)}$ (yellow), $\Lambda^{(2,2)}$ (violet). Assembling re-organized probability matrix $P(Z,K)$ (top right): $P^{(1,1)}(Z,K)$ (red), $P^{(2,1)}(Z,K)$ (green), $P^{(2,2)}(Z,K)$ (violet). Re-organized probability matrix $P(Z,K)$ (bottom left): $P^{(1,1)}(Z,K)$ (red), $P^{(2,1)}(Z,K)$ and $P^{(1,2)}(Z,K)$ (green), $P^{(2,2)}(Z,K)$ (violet). Probability matrix P (bottom right): nodes 1,3,4 are in community 1; nodes 2 and 5 are in community 2.

In reality, K_* and matrices Z_* and P_* are unknown and need to be recovered. If K_* were known, in order to estimate Z_* and P_*, one could permute the rows and the columns of the adjacency matrix A using permutation matrix $\mathscr{P}_{Z,K_*}$ obtaining matrix $A(Z,K_*) = \mathscr{P}^T_{Z,K_*} A \mathscr{P}_{Z,K_*}$ and then, following assumption (3.2), minimize some divergence measure between blocks of $A(Z,K_*)$ and the products $\Lambda^{(k,l)}\,[\Lambda^{(l,k)}]^T$. One of such measures is the Bregman divergence between $A(Z,K_*)$ and $\Lambda^{(k,l)}\,[\Lambda^{(l,k)}]^T$.

The Bregman divergence between vectors x and y associated with a continuously-differentiable, strictly convex function F is defined as

$$D_F(x,y) = F(x) - F(y) - \langle \nabla F(y), x - y \rangle$$

where $\nabla F(y)$ is the gradient of F with respect to y. The Bregman divergence between any matrices X and Y of the same dimension can be defined as the Bregman divergence between their vectorized versions: $D_F(X,Y) = D_F(\text{vec}(X), \text{vec}(Y))$. It is well known that $D_F(X,Y) \geq 0$ for any X and Y and $D_F(X,Y) = 0$ iff $X = Y$. In particular, the Poisson log-likelihood maximization used in Sengupta and Chen (2018) corresponds to minimizing the Bregman divergence with

$$F(x) = \sum_i (x_i \ln x_i - x_i).$$

Under the assumption (3.2) and the constraint (3.7) of Sengupta and Chen (2018), the latter leads to maximization over $\Lambda^{(k,l)}$ and $Z \in \mathcal{M}_{n,K_*}$ of the following quantity

$$l(\Lambda|A) = -D_F(A,\Lambda) = \sum_{k,l=1}^{K_*} \sum_{i=1}^{n_k} \sum_{j=1}^{n_l} \left[A_{i,j}^{(k,l)} \ln \left(\Lambda_i^{(k,l)} \Lambda_j^{(l,k)} \right) - \left(\Lambda_i^{(k,l)} \Lambda_j^{(l,k)} \right) \right]. \tag{3.8}$$

where $A^{(k,l)}$ stands for $A^{(k,l)}(Z,K_*)$, the (k,l)-th block of matrix $A(Z,K_*)$. It is easy to see that the expression (3.8) coincides with the Poisson log-likelihood up to a term which depends on matrix A only, and is independent of P, Z and K_*. Maximization of (3.8) over Λ, under condition (3.7), for given Z and K_*, leads to the estimators of Λ obtained in Sengupta and Chen (2018)

$$\widehat{\Lambda}^{(k,l)} = \frac{A^{(k,l)}(Z,K_*) 1_{n_l}}{\sqrt{1_{n_k}^T A^{(k,l)}(Z,K_*) 1_{n_l}}}; \quad \widehat{\Lambda}^{(l,k)} = \frac{(A^{(k,l)}(Z,K_*))^T 1_{n_k}}{\sqrt{1_{n_k}^T A^{(k,l)}(Z,K_*) 1_{n_l}}}. \tag{3.9}$$

Afterwards, Sengupta and Chen (2018) plug the estimators (3.9) into (3.8), thus, obtaining the likelihood modularity function which they further maximize in order to obtain community assignments.

In the present work, we use the Bregman divergence associated with the Euclidean distance ($F(x) = \|x\|^2$) which, for a given K, leads to the following optimization problem

$$(\widehat{\Lambda}, \hat{Z}) \in \underset{\Lambda, Z}{\operatorname{argmin}} \left\{ \sum_{k,l=1}^{K} \left\| A^{(k,l)}(Z, K) - \Lambda^{(k,l)} [\Lambda^{(l,k)}]^T \right\|_F^2 \right\} \quad \text{s.t.} \quad A(Z, K) = \mathscr{P}_{Z,K}^T A \mathscr{P}_{Z,K}$$

Note that recovery of the components $\Lambda^{(k,l)}$ and $\Lambda^{(l,k)}$ of the products above relies on an identifiability condition of the type (3.7). Since these conditions can be imposed in a variety of ways, we denote $\Theta^{(k,l)} = \Lambda^{(k,l)} [\Lambda^{(l,k)}]^T$ and recover the uniquely defined rank one matrix $\Theta^{(k,l)}$. In addition, since the number of clusters K is unknown, we impose a penalty on K in order to safeguard against choosing too many clusters. Hence, we need to solve the following optimization problem

$$\begin{aligned} (\hat{\Theta}, \hat{Z}, \hat{K}) \in \quad & \underset{\Theta, Z, K}{\operatorname{argmin}} \left\{ \sum_{k,l=1}^{K} \left\| A^{(k,l)}(Z, K) - \Theta^{(k,l)} \right\|_F^2 + \text{Pen}(n, K) \right\} \\ & \text{s.t.} \quad A(Z, K) = \mathscr{P}_{Z,K}^T A \mathscr{P}_{Z,K}, \quad \text{rank}(\Theta^{(k,l)}) = 1; \quad k, l = 1, 2, \cdots, K. \end{aligned} \tag{3.10}$$

Here, $\hat{\Theta}$ is the block matrix with blocks $\hat{\Theta}^{(k,l)}$, $k, l = 1, \dots, \hat{K}$ and $\text{Pen}(n, K)$ will be defined later.

Observe that, if $\hat{Z}$ and $\hat{K}$ were known, the best solution of problem (4.5) would be given by the rank one approximations $\hat{\Theta}^{(k,l)}$ of matrices $A^{(k,l)}(\hat{Z}, \hat{K})$

$$\hat{\Theta}^{(k,l)}(\hat{Z}, \hat{K}) = \Pi_{\hat{u},\hat{v}} \left(A^{(k,l)}(\hat{Z}, \hat{K}) \right) = \hat{\sigma}_1^{(k,l)} \hat{u}^{(k,l)}(\hat{Z}, \hat{K}) (\hat{v}^{(k,l)}(\hat{Z}, \hat{K}))^T, \tag{3.11}$$

where $\hat{\sigma}_1^{(k,l)}$ are the largest singular values of matrices $A^{(k,l)}(\hat{Z}, \hat{K}))$; $\hat{u}^{(k,l)}(\hat{Z}, \hat{K})$, $\hat{v}^{(k,l)}(\hat{Z}, \hat{K})$ are the corresponding singular vectors, and $\Pi_{\hat{u},\hat{v}} \left(A^{(k,l)}(\hat{Z}, \hat{K}) \right)$ is the rank one projection of matrix $A^{(k,l)}(\hat{Z}, \hat{K})$ (see Lemma 4.4.1 for the exact expression). Plugging (4.6) into (4.5), we rewrite optimization problem (4.5) as

$$(\hat{Z}, \hat{K}) \in \underset{Z,K}{\operatorname{argmin}} \left\{ \sum_{k,l=1}^{K} \left\| A^{(k,l)}(Z,K) - \Pi_{\hat{u},\hat{v}} \left(A^{(k,l)}(Z,K) \right) \right\|_F^2 + \operatorname{Pen}(n,K) \right\} \tag{3.12}$$
$$\text{s.t.} \quad A(Z,K) = \mathscr{P}_{Z,K}^T A \mathscr{P}_{Z,K}$$

In order to obtain $(\hat{Z}, \hat{K})$, one needs to solve optimization problem (4.7) for every K, obtaining

$$\hat{Z}_K \in \underset{Z \in \mathcal{M}_{n,K}}{\operatorname{argmin}} \left\{ \sum_{k,l=1}^{K} \left\| A^{(k,l)}(Z,K) - \Pi_{\hat{u},\hat{v}} \left(A^{(k,l)}(Z,K) \right) \right\|_F^2 \right\} \tag{3.13}$$

and then find $\hat{K}$ as

$$\hat{K} \in \underset{K}{\operatorname{argmin}} \left\{ \sum_{k,l=1}^{K} \left\| A^{(k,l)}(\hat{Z}_K,K) - \Pi_{\hat{u},\hat{v}} \left(A^{(k,l)}(\hat{Z}_K,K) \right) \right\|_F^2 + \operatorname{Pen}(n,K) \right\}. \tag{3.14}$$

Note that if the true number of clusters K_* were known, the penalty in (4.5) and (4.7) would be unnecessary.

Remark 1. Advantages of our estimation procedure. There are several advantages of the estimator (4.6) in comparison with estimators (3.9) of Sengupta and Chen (2018). First, rather than obtaining estimators in (3.9) by averaging, we derive the rank one approximations of the unknown sub-matrices of probabilities which lead to the minimal error (see, e.g., Giraud (2015)) even when some of the nodes are misclustered and, therefore, the matrices $P_*^{(k,l)}(\hat{Z}, \hat{K})$ are not necessarily of rank one. Indeed, the estimators obtained by averaging are suboptimal since matrix P_* is contaminated with errors. Second, recoveries of the matrices $\Theta^{(k,l)}$ do not require any identifiability conditions that can be imposed in a variety of ways. Finally, estimators $\widehat{\Lambda}^{(k,k)}$ of vectors $\Lambda^{(k,k)}$ in (3.9) require the knowledge of the diagonal elements of matrix A that are not available. On the contrary, the rank one approximation of a matrix can be achieved in the presence of missing values (see, e.g., Klopp, Lounici and Tsybakov (2017)).

Remark 2. The true community assignment. Sengupta and Chen (2018) show that the likelihood modularity is maximized at the true community assignment provided the, so called, detectability condition holds: for any two distinct communities $\mathcal{N}_l$ and $\mathcal{N}_k$ and any two nodes, $j_1 \in \mathcal{N}_l$ and $j_2 \in \mathcal{N}_k$, the set $\{(P_*)_{i,j_1}/(P_*)_{i,j_2}\}_{i=1}^n$ assumes at least $K_* + 1$ distinct values, where K_* is the true (known) number of clusters and P_* is the unknown true matrix of probabilities. In our case, the correct community assignment is a solution of the optimization problem (4.8) if matrix P_* is a unique combination (up to permutations) of the K^2 rank one matrices. The latter is guaranteed if collections of vectors $\Lambda^{(k,1)}, \ldots, \Lambda^{(k,K_*)}$ are linearly independent for any $k = 1, \ldots, K_*$. Milder conditions can be found in Soltanolkotabi and Candes (2012).

3.6 The Errors of Estimation and Clustering

In this section we evaluate the estimation and the clustering errors. We choose the penalty which, with high probability, exceeds the random errors. In particular, we denote

$$F_1(n, K) = C_1 nK + C_2 K^2 \ln(ne) + C_3(\ln n + n \ln K) \tag{3.15}$$

$$F_2(n, K) = 2 \ln n + 2n \ln K, \tag{3.16}$$

where C_1, C_2 and C_3 are absolute constants. Define the penalty of the form

$$\text{Pen}(n, K) = \left(2 + 16\,\beta_1^{-1}\right) F_1(n, K) + \beta_2^{-1} F_2(n, K), \tag{3.17}$$

where positive parameters β_1 and β_2 are such that $\beta_1 + \beta_2 < 1$. Then, the following statement holds.

Theorem 3.6.1. *Let* $(\hat{\Theta}, \hat{Z}, \hat{K})$ *be a solution of optimization problem* (4.5)*. Construct the estimator*

$\hat{P}$ *of* P_* *of the form*

$$\hat{P} = \mathscr{P}_{\hat{Z},\hat{K}} \hat{\Theta}(\hat{Z}, \hat{K}) \mathscr{P}^T_{\hat{Z},\hat{K}} \tag{3.18}$$

where $\mathscr{P}_{\hat{Z},\hat{K}}$ *is the permutation matrix corresponding to* $(\hat{Z}, \hat{K})$*. Then, for any* $t > 0$ *and* $\tilde{C} = \tilde{C}(C_3)$ *given in* (4.22)*, one has*

$$\mathbb{P}\left\{ \frac{1}{n^2} \left\| \hat{P} - P_* \right\|_F^2 \le \frac{\textit{Pen}(n, K_*)}{(1 - \beta_1 - \beta_2)\, n^2} + \frac{\tilde{C}t}{n^2} \right\} \ge 1 - 3e^{-t}, \tag{3.19}$$

$$\frac{1}{n^2}\, \mathbb{E} \left\| \hat{P} - P_* \right\|_F^2 \le \frac{\textit{Pen}(n, K_*)}{(1 - \beta_1 - \beta_2)\, n^2} + \frac{3\tilde{C}}{n^2} \tag{3.20}$$

Remark 3. The penalty. By rearranging and combining the terms, the penalty in (3.17) can be written in the form

$$\text{Pen}(n, K) = H_1 n K + H_2 K^2 \ln n + H_3 n \ln K, \tag{3.21}$$

where $H_i \equiv H_i(\beta_1, \beta_2, C_1, C_2, C_3)$, $i = 1, 2, 3$, and the estimation errors in (4.20) and (4.21) are proportional to the right hand side of (4.16). The first term in (4.16) corresponds to the error of estimating nK unknown entries of matrix Λ, the second term is associated with estimation of rank K^2 matrix while the last term is due to the clustering of n nodes into K communities. If K grows with n, i.e., $K = K(n) \to \infty$ as $n \to \infty$, then the first term in (4.16) dominates the other two terms. However, in the case of a fixed K, the first and the third terms grow at the same rate as $n \to \infty$. The second term is always of a smaller order provided $K(n)/n \to 0$.

In order to evaluate the clustering error, we assume that the true number of classes $K = K_*$ is known. Let $Z_* \in \mathcal{M}_{n,K_*}$ be the true clustering matrix. Then $\hat{Z} \equiv \hat{Z}_K$ is a solution of the optimization problem (4.8). Note that if Z_* is the true clustering matrix and Z is any other clustering matrix, then the proportion of misclustered nodes can be evaluated as

$$\text{Err}(Z, Z_*) = (2n)^{-1} \min_{\mathscr{P}_K \in \mathcal{P}_K} \|Z\mathscr{P}_K - Z_*\|_1 = (2n)^{-1} \min_{\mathscr{P}_K \in \mathcal{P}_K} \|Z\mathscr{P}_K - Z_*\|_F^2 \tag{3.22}$$

where $\mathcal{P}_K$ is the set of permutation matrices $\mathscr{P}_K : \{1, 2, \cdots, K\} \longrightarrow \{1, 2, \cdots, K\}$. Let

$$\Upsilon(Z_*, \rho) = \left\{ Z \in \mathcal{M}_{n,K} : (2n)^{-1} \min_{\mathscr{P}_K \in \mathcal{P}_K} \|Z\mathscr{P}_K - Z_*\|_1 \geq \rho \right\} \tag{3.23}$$

be the set of clustering matrices with the proportion of misclustered nodes being at least ρ, $\rho < 1$.

The success of clustering in (4.8) relies upon the fact that matrix P_* is a collection of K^2 rank one blocks, so that the operator and the Frobenius norms of each block are the same. On the other hand, if clustering were incorrect, the ranks of the blocks would increase which would lead to the discrepancy between their operator and Frobenius norms. In particular, the following statement is true.

Theorem 3.6.2. *Let $K = K_*$ be the true number of clusters and $Z_* \in \mathcal{M}_{n,K_*}$ be the true clustering matrix. If for some $\alpha_1, \alpha_2, \rho_n \in (0, 1)$, one has*

$$\|P_*\|_F^2 - \frac{1+\alpha_2}{1-\alpha_1} \max_{Z \in \Upsilon(Z_*, \rho_n)} \sum_{k,l=1}^{K} \left\| P_*^{(k,l)}(Z) \right\|_{op}^2 \geq H[C_1 nK + C_2 K^2 \ln(ne) + C_3(n \ln K + t)], \tag{3.24}$$

then, with probability at least $1 - 2e^{-t}$, the proportion of the misclassified nodes is at most ρ_n. Here, $H \equiv H(\alpha_1, \alpha_2)$, is a function of α_1 and α_2 only.

3.7 Supplementary Statements and Proofs

3.7.1 Proof of Theorem 3.6.1.

Denote $\Xi = A - P_*$ and recall that, given matrix P_*, entries $\Xi_{i,j} = A_{i,j} - (P_*)_{ij}$ of Ξ are the independent Bernoulli errors for $1 \le i \le j \le n$ and $A_{i,j} = A_{j,i}$. Then, following notation (3.6), for any Z and K

$$\Xi(Z, K) = \mathscr{P}_{Z,K}^T \Xi \mathscr{P}_{Z,K} \quad \text{and} \quad P_*(Z, K) = \mathscr{P}_{Z,K}^T P_* \mathscr{P}_{Z,K}.$$

Then it follows from (4.5) that

$$\left\| \mathscr{P}_{\hat{Z},\hat{K}}^T A \mathscr{P}_{\hat{Z},\hat{K}} - \hat{\Theta}(\hat{Z}, \hat{K}) \right\|_F^2 + \text{Pen}(n, \hat{K}) \le \left\| \mathscr{P}_{Z_*,K_*}^T A \mathscr{P}_{Z_*,K_*} - \mathscr{P}_{Z_*,K_*}^T P_* \mathscr{P}_{Z_*,K_*} \right\|_F^2 + \text{Pen}(n, K_*)$$

Using the fact that permutation matrices are orthogonal, we can rewrite the previous inequality as

$$\left\| A - \mathscr{P}_{\hat{Z},\hat{K}} \hat{\Theta}(\hat{Z}, \hat{K}) \mathscr{P}_{\hat{Z},\hat{K}}^T \right\|_F^2 + \text{Pen}(n, \hat{K}) \le \| A - P_* \|_F^2 + \text{Pen}(n, K_*). \tag{3.25}$$

Hence, (4.30) and (4.18) yield

$$\left\| A - \hat{P} \right\|_F^2 \le \| A - P_* \|_F^2 + \text{Pen}(n, K_*) - \text{Pen}(n, \hat{K}) \tag{3.26}$$

Subtracting and adding P_* in the norm of the left-hand side of (4.31), we rewrite (4.31) as

$$\left\| \hat{P} - P_* \right\|_F^2 \le \Delta(\hat{Z}, \hat{K}) + \text{Pen}(n, K_*) - \text{Pen}(n, \hat{K}), \tag{3.27}$$

where

$$\Delta(\hat{Z}, \hat{K}) = 2\text{Tr}\left[\Xi^T(\hat{P} - P_*)\right]. \tag{3.28}$$

Again, using orthogonality of the permutation matrices, we can rewrite

$$\Delta(\hat{Z}, \hat{K}) = 2\langle \Xi(\hat{Z}, \hat{K}), (\hat{\Theta}(\hat{Z}, \hat{K}) - P_*(\hat{Z}, \hat{K}))\rangle,$$

where $\langle A, B\rangle = \text{Tr}(A^T B)$. Then, in the block form, $\Delta(\hat{Z}, \hat{K})$ appears as

$$\Delta(\hat{Z}, \hat{K}) = \sum_{k,l=1}^{\hat{K}} \Delta^{(k,l)}(\hat{Z}, \hat{K}) \tag{3.29}$$

where

$$\Delta^{(k,l)}(\hat{Z}, \hat{K}) = 2\left\langle \Xi^{(k,l)}(\hat{Z}, \hat{K}), \Pi_{\hat{u},\hat{v}}\left(A^{(k,l)}(\hat{Z}, \hat{K})\right) - P_*^{(k,l)}(\hat{Z}, \hat{K})\right\rangle$$

and $\Pi_{\hat{u},\hat{v}}$ is defined in (4.63) of Lemma 4.4.1.

Let $\tilde{u} = \tilde{u}^{(k,l)}(\hat{Z}, \hat{K}), \tilde{v} = \tilde{v}^{(k,l)}(\hat{Z}, \hat{K})$ be the singular vectors of $P_*^{(k,l)}(\hat{Z}, \hat{K})$ corresponding to the largest singular value of $P_*^{(k,l)}(\hat{Z}, \hat{K})$. Then, according to Lemma 4.4.1

$$\Pi_{\tilde{u},\tilde{v}}\left(P_*^{(k,l)}(\hat{Z}, \hat{K})\right) = \tilde{u}^{(k,l)}(\hat{Z}, \hat{K})(\tilde{u}^{(k,l)}(\hat{Z}, \hat{K}))^T P_*^{(k,l)}(\hat{Z}, \hat{K})\tilde{v}^{(k,l)}(\hat{Z}, \hat{K})(\tilde{v}^{(k,l)}(\hat{Z}, \hat{K}))^T \tag{3.30}$$

Recall that

$$\Pi_{\hat{u},\hat{v}}(A^{(k,l)}(\hat{Z}, \hat{K})) = \Pi_{\hat{u},\hat{v}}\left[P_*^{(k,l)}(\hat{Z}, \hat{K}) + \Xi^{(k,l)}(\hat{Z}, \hat{K})\right],$$

Then, $\Delta^{(k,l)}(\hat{Z}, \hat{K})$ can be partitioned into the sums of three components

$$\Delta^{(k,l)}(\hat{Z}, \hat{K}) = \Delta_1^{(k,l)}(\hat{Z}, \hat{K}) + \Delta_2^{(k,l)}(\hat{Z}, \hat{K}) + \Delta_3^{(k,l)}(\hat{Z}, \hat{K}), \quad k, l = 1, 2, \cdots, K, \tag{3.31}$$

where

$$\Delta_1^{(k,l)}(\hat{Z},\hat{K}) \;=\; 2\langle \Xi^{(k,l)}(\hat{Z},\hat{K}), \Pi_{\hat{u},\hat{v}}(\Xi^{(k,l)}(\hat{Z},\hat{K}))\rangle \tag{3.32}$$

$$\Delta_2^{(k,l)}(\hat{Z},\hat{K}) \;=\; 2\langle \Xi^{(k,l)}(\hat{Z},\hat{K}), \Pi_{\tilde{u},\tilde{v}}\left(P_*^{(k,l)}(\hat{Z},\hat{K})\right) - P_*^{(k,l)}(\hat{Z},\hat{K})\rangle \tag{3.33}$$

$$\Delta_3^{(k,l)}(\hat{Z},\hat{K}) \;=\; 2\langle \Xi^{(k,l)}(\hat{Z},\hat{K}), \Pi_{\hat{u},\hat{v}}(P_*^{(k,l)}(\hat{Z},\hat{K})) - \Pi_{\tilde{u},\tilde{v}}\left(P_*^{(k,l)}(\hat{Z},\hat{K})\right)\rangle \tag{3.34}$$

With some abuse of notations, for any matrix B, let $\Pi_{\tilde{u},\tilde{v}}\left(B(\hat{Z},\hat{K})\right)$ be the matrix with blocks $\Pi_{\tilde{u},\tilde{v}}\left(B^{(k,l)}(\hat{Z},\hat{K})\right)$, and $\Pi_{\hat{u},\hat{v}}\left(B(\hat{Z},\hat{K})\right)$ be the matrix with blocks $\Pi_{\hat{u},\hat{v}}\left(B^{(k,l)}(\hat{Z},\hat{K})\right)$, $k,l = 1,2,\cdots,\hat{K}$. Then, it follows from (4.34)–(4.37) that

$$\Delta(\hat{Z},\hat{K}) = \Delta_1(\hat{Z},\hat{K}) + \Delta_2(\hat{Z},\hat{K}) + \Delta_3(\hat{Z},\hat{K}) \tag{3.35}$$

where

$$\Delta_1(\hat{Z},\hat{K}) \;=\; 2\langle (\Xi(\hat{Z},\hat{K}), \Pi_{\hat{u},\hat{v}}(\Xi(\hat{Z},\hat{K}))\rangle \tag{3.36}$$

$$\Delta_2(\hat{Z},\hat{K}) \;=\; 2\langle \Xi(\hat{Z},\hat{K}), \Pi_{\tilde{u},\tilde{v}}\left(P_*(\hat{Z},\hat{K})\right) - P_*(\hat{Z},\hat{K})\rangle \tag{3.37}$$

$$\Delta_3(\hat{Z},\hat{K}) \;=\; 2\langle \Xi(\hat{Z},\hat{K}), \Pi_{\hat{u},\hat{v}}(P_*(\hat{Z},\hat{K})) - \Pi_{\tilde{u},\tilde{v}}\left(P_*(\hat{Z},\hat{K})\right)\rangle \tag{3.38}$$

Now, we need to derive an upper bound for each component in (4.34) and (4.38).

Observe that

$$\begin{aligned}\Delta_1^{(k,l)}(\hat{Z},\hat{K}) = 2\langle \Xi^{(k,l)}(\hat{Z},\hat{K}), \Pi_{\hat{u},\hat{v}}(\Xi^{(k,l)}(\hat{Z},\hat{K}))\rangle = 2\left\|\Pi_{\hat{u},\hat{v}}(\Xi^{(k,l)}(\hat{Z},\hat{K}))\right\|_F^2 \\ \le 2\left\|\Xi^{(k,l)}(\hat{Z},\hat{K})\right\|_{op}^2.\end{aligned}$$

Now, fix t and let Ω_1 be the set where $\left\|\Xi(\hat{Z},\hat{K})\right\|_{op}^2 \le F_1(n,\hat{K}) + C_3 t$. According to Lemma 4.4.4,

$$\mathbb{P}(\Omega_1) \geq 1 - \exp(-t), \tag{3.39}$$

and, for $\omega \in \Omega_1$, one has

$$|\Delta_1(\hat{Z}, \hat{K})| \leq 2 \sum_{k,l=1}^{\hat{K}} \left\| \Xi^{(k,l)}(\hat{Z}, \hat{K}) \right\|_{op}^2 \leq 2F_1(n, \hat{K}) + 2C_3 t \tag{3.40}$$

Now, consider $\Delta_2(\hat{Z}, \hat{K})$ given by (4.40). Note that

$$|\Delta_2(\hat{Z}, \hat{K})| = 2\|\Pi_{\tilde{u},\tilde{v}} \left(P_*(\hat{Z}, \hat{K}) \right) - P_*(\hat{Z}, \hat{K})\|_F |\langle \Xi(\hat{Z}, \hat{K}), H_{\tilde{u},\tilde{v}}(\hat{Z}, \hat{K}) \rangle| \tag{3.41}$$

where

$$H_{\tilde{u},\tilde{v}}(\hat{Z}, \hat{K}) = \frac{\Pi_{\tilde{u},\tilde{v}} \left(P_*(\hat{Z}, \hat{K}) \right) - P_*(\hat{Z}, \hat{K})}{\|\Pi_{\tilde{u}\tilde{v}} \left(P_*(\hat{Z}, \hat{K}) \right) - P_*(\hat{Z}, \hat{K})\|_F}$$

Since for any a, b and $\alpha_1 > 0$, one has $2ab \leq \alpha_1 a^2 + b^2/\alpha_1$, obtain

$$|\Delta_2(\hat{Z}, \hat{K})| \leq \alpha_1 \|\Pi_{\tilde{u},\tilde{v}} \left(P_*(\hat{Z}, \hat{K}) \right) - P_*(\hat{Z}, \hat{K})\|_F^2 + 1/\alpha_1 \, |\langle \, \Xi(\hat{Z}, \hat{K}), H_{\tilde{u},\tilde{v}}(\hat{Z}, \hat{K}) \, \rangle|^2 \tag{3.42}$$

Observe that if K and $Z \in \mathcal{M}_{n,K}$ are fixed, then $H_{\tilde{u},\tilde{v}}(Z, K)$ is fixed and, for any K and Z, one has $\|H_{\tilde{u},\tilde{v}}(Z, K)\|_F = 1$. Note also that, for fixed K and Z, permuted matrix $\Xi(Z, K) \in [0, 1]^{n \times n}$ contains independent Bernoulli errors. It is well known that if ξ is a vector of independent Bernoulli errors and h is any fixed vector, then, for any $x > 0$, Hoeffding's inequality yields

$$\mathbb{P}(|\xi^T h|^2 > x) \leq 2 \exp(-x/2)$$

Since $\langle \Xi(Z,K), H_{\tilde{u},\tilde{v}}(Z,K)\rangle = [\text{vec}(\Xi(Z,K))]^T \text{vec}(H_{\tilde{u},\tilde{v}}(Z,K))$, obtain for any fixed K and Z:

$$\mathbb{P}\left(|\langle \Xi(Z,K), H_{\tilde{u},\tilde{v}}(Z,K)\rangle|^2 - x > 0\right) \leq 2\exp(-x/2)$$

Now, applying the union bound, derive

$$\begin{aligned}
&\mathbb{P}\left(|\langle \Xi(\hat{Z},\hat{K}), H_{\tilde{u},\tilde{v}}(\hat{Z},\hat{K})\rangle|^2 - F_2(n,\hat{K}) > 2t\right) \\
\leq &\mathbb{P}\left(\max_{1\leq K\leq n} \max_{Z\in\mathcal{M}_{n,k}} [|\langle \Xi(Z,K), H_{\tilde{u},\tilde{v}}(Z,K)\rangle|^2 - F_2(n,K)] > 2t\right) \\
\leq &2nK^n \exp\left\{-F_2(n,K)/2 - t\right\} = 2\exp(-t),
\end{aligned} \tag{3.43}$$

where $F_2(n,K)$ is defined in (4.28). By Lemma 4.4.2, one has

$$\|\Pi_{\tilde{u},\tilde{v}}\left(P_*(\hat{Z},\hat{K})\right) - P_*(\hat{Z},\hat{K})\|_F^2 \leq \|\Pi_{\hat{u},\hat{v}}\left(P_*(\hat{Z},\hat{K})\right) - P_*(\hat{Z},\hat{K})\|_F^2 \leq \|\hat{P} - P_*\|_F^2.$$

Denote the set on which (4.45) holds by Ω_2^C, so that

$$\mathbb{P}(\Omega_2) \geq 1 - 2\exp(-t). \tag{3.44}$$

Then inequalities (4.44) and (4.45) imply that, for any $\alpha_1 > 0$, $t > 0$ and any $\omega \in \Omega_2$, one has

$$|\Delta_2(\hat{Z},\hat{K})| \leq \alpha_1\|\hat{P} - P_*\|_F^2 + 1/\alpha_1\, F_2(n,\hat{K}) + 2\,t/\alpha_1. \tag{3.45}$$

Now consider $\Delta_3(\hat{Z},\hat{K})$ defined in (4.41) with components (4.37). Note that matrices $\Pi_{\hat{u},\hat{v}}(P_*^{(k,l)}(\hat{Z},\hat{K})) - \Pi_{\tilde{u},\tilde{v}}\left(P_*^{(k,l)}(\hat{Z},\hat{K})\right)$ have rank at most two. Use the fact that (see, e.g., Giraud (2014), page 123)

$$\langle A,B\rangle \leq \|A\|_{(2,r)}\|B\|_{(2,r)} \leq 2\|A\|_{op}\|B\|_F, \quad r = \min\{\text{rank}(A), \text{rank}(B)\}. \tag{3.46}$$

Here $\|A\|_{(2,q)}$ is the Ky-Fan $(2,q)$ norm

$$\|A\|_{(2,q)}^2 = \sum_{j=1}^{q} \sigma_j^2(A) \le \|A\|_F^2,$$

where $\sigma_j(A)$ are the singular values of A. Applying inequality (4.48) with $r = 2$ and taking into account that for any matrix A one has $\|A\|_{(2,2)}^2 \le 2\|A\|_{op}^2$, derive

$$|\Delta_3^{(k,l)}(\hat{Z},\hat{K})| \le 4\|\Xi^{(k,l)}(\hat{Z},\hat{K})\|_{op}\|\Pi_{\hat{u},\hat{v}}(P_*^{(k,l)}(\hat{Z},\hat{K})) - \Pi_{\tilde{u},\tilde{v}}\left(P_*^{(k,l)}(\hat{Z},\hat{K})\right)\|_F.$$

Then, for any $\alpha_2 > 0$, obtain

$$\begin{aligned} |\Delta_3(\hat{Z},\hat{K})| &\le \sum_{k,l=1}^{\hat{K}} |\Delta_3^{(k,l)}(\hat{Z},\hat{K})| \\ &\le \frac{2}{\alpha_2}\sum_{k,l=1}^{\hat{K}} \|\Xi^{(k,l)}(\hat{Z},\hat{K})\|_{op}^2 + 2\alpha_2 \sum_{k,l=1}^{\hat{K}} \|\Pi_{\hat{u},\hat{v}}(P_*^{(k,l)}(\hat{Z},\hat{K})) - \Pi_{\tilde{u},\tilde{v}}\left(P_*^{(k,l)}(\hat{Z},\hat{K})\right)\|_F^2. \end{aligned} \tag{3.47}$$

Note that, by Lemma 4.4.2,

$$\begin{aligned} &\|\Pi_{\hat{u},\hat{v}}(P_*^{(k,l)}(\hat{Z},\hat{K})) - \Pi_{\tilde{u}\tilde{v}}\left(P_*^{(k,l)}(\hat{Z},\hat{K})\right)\|_F^2 \\ \le& 2\|\Pi_{\hat{u},\hat{v}}(P_*^{(k,l)}(\hat{Z},\hat{K})) - P_*^{(k,l)}(\hat{Z},\hat{K})\|_F^2 + 2\|\Pi_{\tilde{u},\tilde{v}}(P_*^{(k,l)}(\hat{Z},\hat{K})) - P_*^{(k,l)}(\hat{Z},\hat{K})\|_F^2 \\ \le& 4\|\Pi_{\hat{u},\hat{v}}(P_*^{(k,l)}(\hat{Z},\hat{K})) - P_*^{(k,l)}(\hat{Z},\hat{K})\|_F^2 \\ \le& 4\|\Pi_{\hat{u},\hat{v}}(A^{(k,l)}(\hat{Z},\hat{K})) - P_*^{(k,l)}(\hat{Z},\hat{K})\|_F^2 = 4\|\hat{\Theta}^{(k,l)}(\hat{Z},\hat{K}) - P_*^{(k,l)}(\hat{Z},\hat{K})\|_F^2 \end{aligned}$$

Therefore,

$$\sum_{k,l=1}^{\hat{K}} \|\Pi_{\hat{u},\hat{v}}(P_*^{(k,l)}(\hat{Z},\hat{K})) - \Pi_{\tilde{u}\tilde{v}}\left(P_*^{(k,l)}(\hat{Z},\hat{K})\right)\|_F^2 \le$$

$$4\left\|\hat{\Theta}(\hat{Z},\hat{K}) - P_*(\hat{Z},\hat{K})\right\|_F^2 = 4\|\hat{P} - P_*\|_F^2 \tag{3.48}$$

Combine inequalities (4.49) and (3.48) and recall that $\left\|\Xi(\hat{Z},\hat{K})\right\|_{op}^2 \le F_1(n,\hat{K}) + C_3\, t$ for $\omega \in \Omega_1$. Then, for any $\alpha_2 > 0$ and $\omega \in \Omega_1$, one has

$$|\Delta_3(\hat{Z},\hat{K})| \le 8\alpha_2\|\hat{P} - P_*\|_F^2 + 2/\alpha_2 F_1(n,\hat{K}) + 2C_3\, t/\alpha_2. \tag{3.49}$$

Now, let $\Omega = \Omega_1 \cap \Omega_2$. Then, (4.42) and (4.46) imply that $\mathbb{P}(\Omega) \ge 1 - 3\exp(-t)$ and, for $\omega \in \Omega$, inequalities (4.43), (4.47) and (4.50) simultaneously hold. Hence, by (4.38), derive that, for any $\omega \in \Omega$,

$$|\Delta(\hat{Z},\hat{K})| \le (2+2/\alpha_2)F_1(n,\hat{K})) + 1/\alpha_1\, F_2(n,\hat{K}) + (\alpha_1+8\alpha_2)\|\hat{P}-P_*\|_F^2 + 2(C_3+1/\alpha_1+C_3/\alpha_2)\, t.$$

Combination of the last inequality and (3.27) yields that, for $\alpha_1 + 8\alpha_2 < 1$ and any $\omega \in \Omega$,

$$(1-\alpha_1-8\alpha_2)\left\|\hat{P} - P_*\right\|_F^2 \le \left(2 + \frac{2}{\alpha_2}\right) F_1(n,\hat{K}) + \frac{1}{\alpha_1}F_2(n,\hat{K}) + \text{Pen}(n,K_*) - \text{Pen}(n,\hat{K}) + 2(C_3+1/\alpha_1+C_3/\alpha_2)$$

Setting $\text{Pen}(n,K) = (2+2/\alpha_2)F_1(n,K) + 1/\alpha_1 F_2(n,K)$ and dividing by $(1-\alpha_1-8\alpha_2)$, obtain that

$$\mathbb{P}\left\{\|\hat{P} - P_*\|_F^2 \le (1-\alpha_1-8\alpha_2)^{-1}\,\text{Pen}(n,K_*) + \tilde{C}\, t\right\} \ge 1 - 3e^{-t} \tag{3.50}$$

where

$$\tilde{C} = 2\,(1-\alpha_1-8\alpha_2)^{-1}\,(C_3 + 1/\alpha_1 + C_3/\alpha_2) \tag{3.51}$$

In order to derive (4.20), set $\beta_1 = 8\alpha_2$ and $\beta_2 = \alpha_1$. In order to obtain the upper bound (4.21) note that for $\xi = \|\hat{P} - P_*\|_F^2 - (1 - \beta_1 - \beta_2)^{-1} \operatorname{Pen}(n, K_*)$, one has $\mathbb{E}\|\hat{P} - P_*\|_F^2 = (1 - \beta_1 - \beta_2)^{-1} \operatorname{Pen}(n, K_*) + \mathbb{E}\xi$, where

$$\mathbb{E}\xi \leq \int_0^\infty \mathbb{P}(\xi > z) dz = \tilde{C} \int_0^\infty \mathbb{P}(\xi > \tilde{C} t) dt \leq \tilde{C} \int_0^\infty 3\, e^{-t}\, dt = 3\tilde{C},$$

which yields (4.21).

3.7.2 Proof of Theorem 3.6.2.

Since we have

$$\hat{Z}_K \in \underset{Z \in \mathcal{M}_{n,K}}{\operatorname{argmin}} \left\{ \sum_{k,l=1}^{K} \left\| A^{(k,l)}(Z, K) - \Pi_{\hat{u},\hat{v}} \left(A^{(k,l)}(Z, K) \right) \right\|_F^2 \right\}$$

so that

$$\sum_{k,l=1}^{K} \left\| A^{(k,l)}(\hat{Z}) - \Pi_{\hat{u},\hat{v}} \left(A^{(k,l)}(\hat{Z}) \right) \right\|_F^2 \leq \sum_{k,l=1}^{K} \left\| A^{(k,l)}(Z_*) - \Pi_{\hat{u},\hat{v}} \left(A^{(k,l)}(Z_*) \right) \right\|_F^2 \tag{3.52}$$

Observe that for any $Z \in \mathcal{M}_{n,K}$, one has

$$\sum_{k,l=1}^{K} \left\| A^{(k,l)}(Z) - \Pi_{\hat{u},\hat{v}} \left(A^{(k,l)}(Z) \right) \right\|_F^2 = \sum_{k,l=1}^{K} \left\{ \left\| A^{(k,l)}(Z) \right\|_F^2 - \left\| \Pi_{\hat{u},\hat{v}} \left(A^{(k,l)}(Z) \right) \right\|_F^2 \right\},$$

so that, due to $\sum_{k,l=1}^{K} \left\| A^{(k,l)}(Z) \right\|_F^2 = \|A\|_F^2$, (4.53) can be re-written as

$$\sum_{k,l=1}^{K} \left\| \Pi_{\hat{u},\hat{v}} \left(A^{(k,l)}(\hat{Z}) \right) \right\|_F^2 \geq \sum_{k,l=1}^{K} \left\| \Pi_{\hat{u},\hat{v}} \left(A^{(k,l)}(Z_*) \right) \right\|_F^2 \tag{3.53}$$

Applying Proposition 6.2 of Giraud (2015), obtain

$$\|\Pi_{\hat{u},\hat{v}}(A^{(k,l)}(Z)) - P_*^{(k,l)}(Z)\|_F^2 \le \frac{(2+\theta)^2}{\theta^2} \sum_{r=2}^{\min\{n_k,n_l\}} \sigma_r^2 P_*^{(k,l)}(Z) + \frac{2(1+\theta)(2+\theta)}{\theta} \|\Xi^{(k,l)}(Z)\|_{op}^2,$$

where $\theta > 0$ is an arbitrary constant, P_* is the true matrix of probabilities, $\Xi^{(k,l)}(Z) = A^{(k,l)}(Z) - P_*^{(k,l)}(Z)$, and $\sigma_r(B)$ is the $r-$th largest singular value of B. Since matrix $P_*^{(k,l)}(Z_*)$ has rank one, the previous inequality yields for $\theta = \sqrt{2}$

$$\|\Pi_{\hat{u},\hat{v}}\left(A^{(k,l)}(Z_*)\right) - P_*^{(k,l)}(Z_*)\|_F^2 \le 2(1+\sqrt{2})^2 \|\Xi^{(k,l)}(Z_*)\|_{op}^2 \tag{3.54}$$

Using Lemma 4.4.3, derive for any $t > 0$ that

$$\mathbb{P}\left\{\sum_{k,l=1}^{K} \left\|\Xi^{(k,l)}(Z_*)\right\|_{op}^2 \le C_1 nK + C_2 K^2 \ln(ne) + C_3\, t\right\} \ge 1 - \exp(-t). \tag{3.55}$$

Also, since $\text{card}(\mathcal{M}_{n,K}) = K^n$, replacing t by $n \ln K + t$ and applying union bound, obtain

$$\mathbb{P}\left\{\sum_{k,l=1}^{K} \|\Xi^{(k,l)}(\hat{Z})\|_{op}^2 \le C_1 nK + C_2 K^2 \ln(ne) + C_3(n \ln K + t)\right\} \ge 1 - \exp(-t). \tag{3.56}$$

Note that for any $\alpha_1 \in (0,1)$, using the inequality $(a+b)^2 \ge (1-\alpha_1)a^2 + (1-\alpha_1^{-1})b^2$, obtain

$$\left\|\Pi_{\hat{u},\hat{v}}\left(A^{(k,l)}(Z_*)\right)\right\|_F^2 = \left\|\Pi_{\hat{u},\hat{v}}\left(A^{(k,l)}(Z_*)\right) - P_*^{(k,l)}(Z_*) + P_*^{(k,l)}(Z_*)\right\|_F^2 \ge$$
$$(1-\alpha_1)\left\|P_*^{(k,l)}(Z_*)\right\|_F^2 - \left(\alpha_1^{-1} - 1\right)\left\|\Pi_{\hat{u},\hat{v}}\left(A^{(k,l)}(Z_*)\right) - P_*^{(k,l)}(Z_*)\right\|_F^2$$

Combining the last inequality with (3.54) and taking a sum, obtain

$$\sum_{k,l=1}^{K} \left\| \Pi_{\hat{u},\hat{v}} \left(A^{(k,l)}(Z_*) \right) \right\|_F^2 \geq (1-\alpha_1) \|P_*\|_F^2 - 2(1+\sqrt{2})^2 \left(\frac{1}{\alpha_1} - 1 \right) \sum_{k,l=1}^{K} \left\| \Xi^{(k,l)}(Z_*) \right\|_{op}^2, \quad (3.57)$$

where we used the fact that $\left\| P_*(Z_*) \right\|_F = \|P_*\|_F$. On the other hand, for any $Z \in \mathcal{M}_{n,K}$ and any $\alpha_2 > 0$, using the inequality $(a+b)^2 \leq (1+\alpha_2)a^2 + (1+\alpha_2^{-1})b^2$, obtain

$$\left\| \Pi_{\hat{u},\hat{v}} \left(A^{(k,l)}(Z) \right) \right\|_F^2 \leq (1+\alpha_2) \left\| \Pi_{\hat{u},\hat{v}} \left(P_*^{(k,l)}(Z) \right) \right\|_F^2 + \left(1 + \alpha_2^{-1} \right) \left\| \Pi_{\hat{u},\hat{v}} \left(\Xi^{(k,l)}(Z) \right) \right\|_F^2,$$

so that

$$\left\| \Pi_{\hat{u},\hat{v}} \left(A^{(k,l)}(Z) \right) \right\|_F^2 \leq (1+\alpha_2) \left\| P_*^{(k,l)}(Z) \right\|_{op}^2 + \left(1 + \alpha_2^{-1} \right) \left\| \Xi^{(k,l)}(Z) \right\|_{op}^2 \quad (3.58)$$

Now, we prove the theorem by contradiction. Assume that $\hat{Z} \in \Upsilon(Z_*, \rho_n)$ is the solution of optimization problem (4.8). Then, inequality (3.53) holds. Combining (3.53), (3.57) and (3.58), obtain that

$$(1-\alpha_1)\|P_*\|_F^2 - 2(1+\sqrt{2})^2 \left(\frac{1}{\alpha_1} - 1 \right) \sum_{k,l=1}^{K} \left\| \Xi^{(k,l)}(Z_*) \right\|_{op}^2 \leq$$
$$(1+\alpha_2) \sum_{k,l=1}^{K} \left\| P_*^{(k,l)}(\hat{Z}) \right\|_{op}^2 + \left(1 + \alpha_2^{-1} \right) \sum_{k,l=1}^{K} \left\| \Xi^{(k,l)}(\hat{Z}) \right\|_{op}^2$$

Due to (4.55) and (3.56), with probability at least $1 - 2\exp(-t)$, the last inequality yields

$$(1-\alpha_1)\, \|P_*\|_F^2 - (1+\alpha_2) \sum_{k,l=1}^{K} \|P_*^{(k,l)}(\hat{Z})\|_{op}^2 \leq (1-\alpha_1)\, H[C_1 nK + C_2 K^2 \ln(ne) + C_3(n \ln K + t)],$$

where

$$H \equiv H(\alpha_1, \alpha_2) = \frac{2(1+\sqrt{2})^2}{\alpha_1} + \frac{1+\alpha_2}{\alpha_2(1-\alpha_1)}. \tag{3.59}$$

The latter contradicts (3.24), since $\hat{Z} \in \Upsilon(Z_*, \rho_n)$, which completes the proof.

3.7.3 Supplementary Lemmas and Proofs

Lemma 3.7.1. *For any matrices $A, B \in \mathbb{R}^{m\times n}$ and any unit vectors $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$, let*

$$\Pi_{u,v}(A) = (uu^T)A(vv^T) \tag{3.60}$$

denote the projection of matrix A on the vectors (u, v). Then,

$$\langle \Pi_{u,v}(B), A - \Pi_{u,v}(A)\rangle = 0. \tag{3.61}$$

Furthermore, if we let $\hat{u}$ and $\hat{v}$ be the singular vectors of matrix A corresponding to its largest singular value σ, the best rank one approximation of A is given by

$$\Pi_{\hat{u},\hat{v}}(A) = (\hat{u}\hat{u}^T)A(\hat{v}\hat{v}^T) = \sigma\hat{u}\hat{v}^T. \tag{3.62}$$

Lemma 3.7.2. *Let $A = P + \Xi$. Denote by $(\hat{u}, \hat{v})$ and (u, v) the pairs of singular vectors of matrices A and P, respectively, corresponding to the largest singular values. Then,*

$$\|\Pi_{u,v}(P) - P\|_F \le \|\Pi_{\hat{u},\hat{v}}(P) - P\|_F \le \|\Pi_{\hat{u},\hat{v}}(A) - P\|_F \tag{3.63}$$

where $\Pi_{u,v}(\cdot)$ is defined in (4.63).

Proof. The first inequality in (4.64) is true because $\Pi_{u,v}(P)$ is the best rank one approximation of P. Validity of the second inequality in (4.64) follows from

$$\|\Pi_{\hat{u},\hat{v}}(A) - P\|_F^2 = \|\Pi_{\hat{u},\hat{v}}(P) - P + \Pi_{\hat{u},\hat{v}}(\Xi)\|_F^2 = \|\Pi_{\hat{u},\hat{v}}(P) - P\|_F^2 + \|\Pi_{\hat{u},\hat{v}}(\Xi)\|_F^2$$

Lemma 3.7.3. *, Let elements of matrix $\Xi \in (-1,1)^{n\times n}$ be independent Bernoulli errors. Let matrix Ξ be partitioned into K^2 sub-matrices $\Xi^{(k,l)}$, $k,l = 1,\cdots,K$. Then, for any $x > 0$*

$$\mathbb{P}\left\{\sum_{k,l=1}^{K} \left\|\Xi^{(k,l)}\right\|_{op}^2 \le C_1 nK + C_2 K^2 \ln(ne) + C_3 x\right\} \ge 1 - \exp(-x), \tag{3.64}$$

where C_1, C_2 and C_3 are absolute constants independent of n and K.

Proof. Consider vectors ξ and μ with elements $\xi_{k,l} = \|\Xi^{(k,l)}\|_{op}$ and $\mu_{k,l} = \mathbb{E}\|\Xi^{(k,l)}\|_{op}$, $k,l = 1,\cdots,K$, and let $\eta = \xi - \mu$. Then,

$$\Delta = \sum_{k,l=1}^{K} \left\|\Xi^{(k,l)}\right\|_{op}^2 = \|\xi\|^2 \le 2\|\eta\|^2 + 2\|\mu\|^2 \tag{3.65}$$

Hence, we need to construct the upper bounds for $\|\eta\|^2$ and $\|\mu\|^2$.

We start with constructing upper bounds for $\|\mu\|^2$. Let $\Xi_{i,j}^{(k,l)}$ be elements of the $(n_k \times n_l)$-dimensional matrix $\Xi^{(k,l)}$. Then, $\mathbb{E}(\Xi_{i,j}^{(k,l)}) = 0$ and, by Hoeffding's inequality, $\mathbb{E}\left\{\exp(\lambda \Xi_{i,j}^{(k,l)})\right\} \le \exp\left(\lambda^2/8\right)$. Taking into account that Bernoulli errors are bounded by one in absolute value and applying Corollary 3.3 of Bandeira and van Handel (2016) with $m = n_k$, $n = n_l$, $\sigma_* = 1$, $\sigma_1 = \sqrt{n_l}$ and $\sigma_2 = \sqrt{n_k}$, obtain

$$\mu_{k,l} \le C_0(\sqrt{n_k} + \sqrt{n_l} + \sqrt{\ln(n_k \wedge n_l)})$$

where C_0 is an absolute constant independent of n_k and n_l. Therefore,

$$\|\mu\|^2 \le 3C_0^2 \sum_{k,l=1}^{K} (n_k + n_l + \ln(n_k \wedge n_l)) \le 6C_0^2 nK + 3C_0^2 K^2 \ln n. \tag{3.66}$$

Next, we show that, for any fixed partition, $\eta_{k,l} = \xi_{k,l} - \mu_{k,l}$ are independent sub-gaussian random variables when $1 \le k \le l \le K$. Independence follows from the conditions of Lemma 4.4.3. To prove the sub-gaussian property, use Talagrand's concentration inequality (Theorem 6.10 of Boucheron et al. (2013)): if $\Xi_1, \Xi_2, \Xi_3, \cdots, \Xi_n$ are independent random variables taking values in the interval $[0,1]$ and $f : [0,1]^n \to R$ is a separately convex function such that $|f(x) - f(y)| \le \|x - y\|$ for all $x, y \in [0,1]^n$, then, for $Z = f(\Xi_1, \Xi_2, \Xi_3, \cdots, \Xi_n)$ and any $t > 0$, one has

$$\mathbb{P}(Z > \mathbb{E}Z + t) \le \exp(-t^2/2). \tag{3.67}$$

Apply this theorem to vectors $\zeta_{k,l} = \text{vec}(\Xi^{(k,l)}) \in [0,1]^{n_k \times n_l}$ and $f(\Xi^{(k,l)}) = f(\zeta_{k,l}) = \left\|\Xi^{(k,l)}\right\|_{op}$. Note that, for any two matrices Ξ and $\tilde{\Xi}$ of the same size, one has $\|\Xi - \tilde{\Xi}\|_{op}^2 \le \|\Xi - \tilde{\Xi}\|_F^2 = \|\text{vec}(\Xi) - \text{vec}(\tilde{\Xi})\|^2$. Then, applying Talagrand's inequality with $Z = \|\Xi^{(k,l)}\|_{op}$ and $Z = -\|\Xi^{(k,l)}\|_{op}$, obtain

$$\mathbb{P}\left(\left|\|\Xi^{(k,l)}\|_{op} - \mathbb{E}\|\Xi^{(k,l)}\|_{op}\right| > t\right) \le 2\exp(-t^2/2).$$

Now, use the Lemma 5.5 of Vershynin (2012) which states that the latter implies that for any $t > 0$ and some absolute constant $C_4 > 0$,

$$\mathbb{E}\left[\exp(t\eta_{k,l})\right] = \mathbb{E}\left[\exp(t(\xi_{k,l} - \mu_{k,l}))\right] \le \exp(C_4 t^2/2), \;\; C_4 > 0. \tag{3.68}$$

Hence, $\eta_{k,l}$ are independent sub-gaussian random variables when $1 \le k \le l \le K$.

Now, we obtain an upper bound for $\|\eta\|^2$. Use Theorem 2.1 of Hsu et al. (2012) which states that for any matrix A, if for some $\sigma > 0$ and any vector h one has $\mathbb{E}[\exp(h^T \tilde{\eta})] \le \exp(\|h\|^2 \sigma^2/2)$, then, for any $x > 0$,

$$\mathbb{P}\left\{\|A\tilde{\eta}\|^2 \ge \sigma^2(\mathrm{Tr}(A^T A) + 2\sqrt{\mathrm{Tr}((A^T A)^2)\, x} + 2\|A^T A\|_{op}\, x)\right\} \le \exp(-x). \tag{3.69}$$

Applying (3.69) with $A = I_{K(K+1)/2}$ and $\sigma^2 = C_4$ to a sub-vector $\tilde{\eta}$ of η which contains components $\eta_{k,l}$ with $1 \le k \le l \le K$, obtain

$$\mathbb{P}\left\{\|\tilde{\eta}\|^2 \ge C_4\left(K(K+1)/2 + \sqrt{2\, K(K+1)\, x} + 2x\right)\right\} \le \exp(-x).$$

Since $\|\eta\|^2 \le 2\|\tilde{\eta}\|^2$, derive

$$\mathbb{P}\left\{\|\eta\|^2 \ge 2C_4 K(K+1) + 6C_4 x\right\} \le \exp\left(-x\right) \tag{3.70}$$

Combination of formulas (4.66) and (4.69) yield

$$\mathbb{P}\left\{\|\xi\|^2 \le 2\|\mu\|^2 + 4C_4 K(K+1) + 12C_4 x\right\} \ge 1 - \exp\left(-x\right)$$

Plugging in $\|\mu\|^2$ from (4.67) into the last inequality, derive for any $x > 0$ that

$$\mathbb{P}\left\{\|\xi\|^2 \le 12C_0^2 nK + 6C_0^2 K^2 \ln n + 4C_4 K(K+1) + 12C_4 x\right\} \ge 1 - \exp\left(-x\right). \tag{3.71}$$

Since $K(K+1) \le 2K^2$ and $6C_0^2 K^2 \ln n + 8C_4 K^2 \le \max(6C_0^2, 8C_4) K^2 \ln(ne)$, inequality (4.65) holds with $C_1 = 12C_0^2$, $C_2 = \max(6C_0^2, 8C_4)$ and $C_3 = 12C_4$.

□

Lemma 3.7.4. *For any* $t > 0$*,*

$$\mathbb{P}\left\{\sum_{k,l=1}^{\hat{K}} \left\|\Xi^{(k,l)}(\hat{Z},\hat{K})\right\|_{op}^{2} - F_1(n,\hat{K}) \le C_3 t\right\} \ge 1 - \exp(-t). \tag{3.72}$$

where $F_1(n, K)$ *is given by* (4.26).

Proof. Using Lemma 4.4.3, for any fixed K and $Z \in \mathcal{M}_{n,K}$, obtain

$$\mathbb{P}\left\{\sum_{k,l=1}^{K} \|\Xi^{(k,l)}(Z,K)\|_{op}^{2} - C_1 nK - C_2 K^2 \ln(ne) - C_3 x \ge 0\right\} \le \exp(-x).$$

Application of the union bound over $Z \in \mathcal{M}_{n,K}$ and $K \in [1, n]$ and setting $x = t + \ln n + n \ln K$ yields

$$\begin{aligned}
&\mathbb{P}\left\{\sum_{k,l=1}^{\hat{K}} \|\Xi^{(k,l)}(\hat{Z},\hat{K})\|_{op}^{2} - C_1 n\hat{K} - C_2\hat{K}^2 \ln(ne) - C_3 t - C_3 \ln n - C_3 n \ln \hat{K} \ge 0\right\}\\
\le&\mathbb{P}\left\{\max_{1\le K\le n}\max_{Z\in\mathcal{M}_{n,K}}\left(\sum_{k,l=1}^{K}\|\Xi^{(k,l)}(Z,K)\|_{op}^{2} - F_1(n,K)\right) \ge C_3 t\right\}\\
\le&\sum_{k=1}^{n}\sum_{Z\in\mathcal{M}_{n,K}}\mathbb{P}\left\{\sum_{k,l=1}^{K}\left\|\Xi^{(k,l)}(Z,K)\right\|_{op}^{2} - F_1(n,K) \ge C_3 t\right\}\\
\le& nK^n \exp\{-t - \ln n - n\ln K\} = \exp(-t),
\end{aligned}$$

which completes the proof. □

CHAPTER 4: ESTIMATION AND CLUSTERING IN SPARSE PABM

4.1 Sparsity in Block Models

The real life networks are usually sparse in a sense that a large number of nodes have small degrees. One of the shortcomings of both the SBM and the DCBM is that they do not allow to efficiently model sparsity in networks. Indeed, for the SBM, it is not realistic to assume that all nodes in a pair of communities have no connections, hence, in the SBM setting, one does not assume that the block probabilities $B_{k,l} = 0$ for some k and l. The DCBM is not very different in this respect since setting any node-specific weight to zero will force the respective node to be totally disconnected from the network. For this reason, unlike in other numerous statistical settings, sparsity in block models is defined as a low maximum probability of connections between the nodes: $\max_{i,j} P_{i,j} \leq \rho(n)$ where $\rho(n) \rightarrow 0$ as $n \rightarrow \infty$ (see, e.g., Klopp, Lounici and Tsybakov (2017), Lei and Rinaldo (2015)). As a result, high degree nodes become very unlikely. In addition to being unrealistic, the above definition of sparsity has other drawbacks. In particular, one has to estimate *every* probability of connections $B_{k,l}$, no matter how small it is, and, in many settings (see, e.g., Klopp, Lounici and Tsybakov (2017)), in order to take advantage of the fact that $P_{i,j}$ are bounded above by $\rho(n)$, one needs to incorporate this unknown value into the estimation process.

On the contrary, the PABM setting allows some connection probabilities to be zero while keeping average connection probabilities between classes above certain level and the network connected. This is possible only in the PABM context due to the flexible modeling of connection probabilities. The idea of setting some infinitesimally small probabilities of connections to zero is quite attractive. Indeed, it is well known that, when many of the elements of a vector or a matrix are identical zeros, identifying those zeros and estimating the rest of the elements leads to a smaller error than when this information is ignored. Similarly, allowing structural sparsity (i.e., setting connection probabilities to zero rather than to a very small positive number) not only leads to better understanding of network topology but leads to more precise estimation of the probability matrix P_*.

In the context of PABM, setting $\Lambda_i^{(k,l)} = 0$ simply means that that node i in class k is not active ("popular") in class l. This, nevertheless, does not prevent this node from having high probability of connection with nodes in another class. Setting some elements of vectors $\Lambda^{(k,l)}$ to zero will merely lead to some of the rows (columns) of sub-matrices $P^{(k,l)}(Z, K)$ being zero. Moreover, since $A_{i,j}$ are Bernoulli variables with the means $P_{i,j}$, those zeros are fairly easy to identify since $P_{i,j} = 0$ leads to $A_{i,j} = 0$.

4.2 Estimation and Clustering in Sparse PABM

In this section, we consider the problem of estimation and clustering of the true matrix P_* of the probabilities of the connection between the nodes.

4.2.1 The Structure of the Probability Matrix

Consider block $P_*^{(k,l)}(Z_*, K_*)$ of the rearranged version $P_*(Z_*, K_*)$ of P_*. Let $\Lambda_* \equiv \Lambda(Z_*, K_*) \in [0,1]^{n \times K_*}$ be a block matrix with each column l partitioned into K_* blocks $\Lambda_*^{(k,l)} \equiv \Lambda_*^{(k,l)}(Z_*, K_*)$. Here, $\Lambda_*^{(k,l)} \in [0,1]^{n_k}$ and $\Lambda_*^{(l,k)} \in [0,1]^{n_l}$ are the column vectors and $P_*^{(k,l)}(Z_*, K_*)$ follows (3.2), i.e., $P_*^{(k,l)}(Z_*, K_*) = \Lambda_*^{(k,l)} [\Lambda_*^{(l,k)}]^T$. Hence, $P_*^{(k,l)}(Z_*, K_*)$ are rank-one matrices such that $P_*^{(k,l)}(Z_*, K_*) = [P_*^{(l,k)}(Z_*, K_*)]^T$ and that each pair of blocks $P_*^{(k,l)}$ and $P_*^{(l,k)}$, involves a unique combination of vectors $\Lambda_*^{(k,l)}$ and $\Lambda_*^{(l,k)}$, $k, l = 1, \dots, K_*$.

Vectors $\Lambda_*^{(k,l)}$ and $\Lambda_*^{(l,k)}$ describe the heterogeneity of the connections of nodes in the pair of communities (k,l). While, on the average, those communities can be connected, some nodes in community k may have no interaction with nodes in community l or vice versa, so that some of the elements of vectors $\Lambda_*^{(k,l)}$ and $\Lambda_*^{(l,k)}$ can be identical zeros. Denote by $J_* \equiv J_*(Z_*, K_*) = \bigcup_{k,l=1}^{K} (J_*)_{k,l}$ the set of indices of all nonzero elements of matrix Λ_*, where

$$(J_*)_{k,l} \equiv (J_*)_{k,l}(Z_*, K_*) = \{i: \ (\Lambda_*)_i^{(k,l)} \neq 0\}, \quad J_*^{(k,l)} = (J_*)_{k,l} \times (J_*)_{l,k}, \tag{4.1}$$

are, respectively, the true support of vector $\Lambda_*^{(k,l)}$ and the set of all ordered pairs of indices (positions) of non-zero elements of sub-matrix $P_*^{(k,l)}(Z_*, K_*)$. Here, the elements of $(J_*)_{k,l}$ are enumerated by their corresponding rows in matrix Λ_*. Then,

$$(P_*)_{i,j}^{(k,l)}(Z_*, K_*) > 0 \quad \text{iff} \quad (i,j) \in J_*^{(k,l)}$$

and row i and column j of $P_*^{(k,l)}(Z_*, K_*)$ are equal to zero if $i \notin (J_*)_{k,l}$ or $j \notin (J_*)_{l,k}$.

Note that the set $J_* \equiv J_*(Z_*, K_*)$ relies upon the true clustering defined by K_* and Z_*. One can also consider sparsity sets $(\breve{J}_*)_{k,l} \equiv (\breve{J}_*)_{k,l}(Z, K)$ and $\breve{J}_{k,l} \equiv \breve{J}_{k,l}(Z, K)$ for an arbitrary K and matrix $Z \in \mathcal{M}_{n,K}$

$$(\breve{J}_*)_{k,l} = \{i: \ (P_*)_{i,j}^{(k,l)}(Z, K) \neq 0, \ j = 1, \ldots, n_l\}, \ \breve{J}_{k,l} = \{i: \ A_{i,j}^{(k,l)}(Z, K) \neq 0, \ j = 1, \ldots, n_l\}, \tag{4.2}$$

where the elements of $(\breve{J}_*)_{k,l}$ and $\breve{J}_{k,l}$ are enumerated by their corresponding rows in matrices P_* and A, respectively. Examples of the sets $(J_*)_{k,l}$, $(J_*)^{(k,l)}$, $(\breve{J}_*)_{k,l}$ and $(\breve{J}_*)^{k,l}$ are considered in Section 4.2.3.

For any sparsity sets $J_{k,l} \equiv J_{k,l}(Z, K)$, define, similarly to (4.1),

$$J = \bigcup_{k,l=1}^{K} J_{k,l} \quad \text{with} \quad J^{(k,l)} = J_{k,l} \times J_{l,k} \tag{4.3}$$

It follows from the definitions (4.2) and (4.3) that for any K, $Z \in \mathcal{M}_{n,K}$ and $k, l = 1, \dots, K$

$$\breve{J}_{k,l}(Z, K) \subseteq (\breve{J}_*)_{k,l}(Z, K) \quad \text{and} \quad \breve{J}(Z, K) \subseteq \breve{J}_*(Z, K). \tag{4.4}$$

4.2.2 Optimization Procedure for Estimation and Clustering

Observe that although matrices $P_*^{(k,l)}(Z_*, K_*)$ and the sets $J_*^{(k,l)}$ are well defined, vectors $\Lambda_*^{(k,l)}$ and $\Lambda_*^{(l,k)}$ can be determined only up to a multiplicative constant. In order to avoid this ambiguity, denote $\Theta_*^{(k,l)} = \Lambda_*^{(k,l)}[\Lambda_*^{(l,k)}]^T$ and recover matrix Θ_* with the uniquely defined rank one blocks $\Theta_*^{(k,l)}$ and their supports $J_*^{(k,l)}$, $k, l = 1, \dots, K_*$. Then, one needs to solve the following optimization problem

$$\begin{aligned}
(\hat{\Theta}, \hat{Z}, \hat{J}, \hat{K}) \in \ & \underset{\Theta, Z, J, K}{\text{argmin}} \left\{ \sum_{k,l=1}^{K} \left\| A^{(k,l)}(Z, K) - \Theta^{(k,l)}(Z, J, K) \right\|_F^2 + \text{Pen}(n, J, K) \right\} \\
& \text{s.t.} \quad A(Z, K) = \mathscr{P}_{Z,K}^T A \mathscr{P}_{Z,K}, \ Z \in \mathcal{M}_{n,K}, \\
& \text{supp}(\Theta^{(k,l)}) = J^{(k,l)} = J_{k,l} \times J_{l,k}, \quad \text{rank}(\Theta^{(k,l)}) = 1, \quad k, l = 1, 2, \cdots, K.
\end{aligned} \tag{4.5}$$

Here, $\hat{\Theta}$ is the block matrix with blocks $\hat{\Theta}^{(k,l)}$, $k, l = 1, \dots, K$.

Observe that, if $\hat{Z}$, $\hat{J}$ and $\hat{K}$ were known, the best solution of problem (4.5) would be given by the best rank one approximations $\hat{\Theta}^{(k,l)}$ of matrices $A^{(k,l)}(\hat{Z}, \hat{K})$ restricted to the sets $\hat{J}^{(k,l)}$ of indices of nonzero elements:

$$\hat{\Theta}^{(k,l)}(\hat{Z},\hat{J},\hat{K}) = \Pi_{(1)}\left(\Pi_{\hat{J}^{(k,l)}}(A^{(k,l)}(\hat{Z},\hat{K}))\right), \tag{4.6}$$

where $\Pi_{J^{(k,l)}}\left(A^{(k,l)}\right)$ is the projection of matrix $A^{(k,l)}$ onto the set of matrices with the support $J^{(k,l)}$ and $\Pi_{(1)}$ is the best rank one approximation of a matrix. Plugging (4.6) into (4.5), we rewrite optimization problem (4.5) as

$$(\hat{Z},\hat{J},\hat{K}) \in \underset{Z,J,K}{\operatorname{argmin}}\left\{\sum_{k,l=1}^{K}\|A^{(k,l)}(Z,K) - \Pi_{(1)}[\Pi_{J^{(k,l)}}(A^{(k,l)}(Z,K))]\|_F^2 + \text{Pen}(n,J,K)\right\} \tag{4.7}$$

$$\text{s.t.} \quad A(Z,K) = \mathscr{P}_{Z,K}^T A \mathscr{P}_{Z,K}, \quad Z \in \mathcal{M}_{n,K},$$
$$J^{(k,l)} \equiv J^{(k,l)}(Z,K) = J_{k,l}(Z,K) \times J_{l,k}(Z,K).$$

In practice, in order to obtain $(\hat{Z},\hat{J},\hat{K})$, one needs to solve optimization problem (4.7) for every K, obtaining

$$(\hat{Z}_K,\hat{J}_K) \in \underset{Z,J}{\operatorname{argmin}}\left\{\sum_{k,l=1}^{K}\left\|A^{(k,l)}(Z,K) - \Pi_{(1)}\left(\Pi_{J^{(k,l)}}(A^{(k,l)}(Z,K))\right)\right\|_F^2 + \text{Pen}(n,J,K)\right\} \tag{4.8}$$

$$\text{s.t.} \quad A(Z,K) = \mathscr{P}_{Z,K}^T A \mathscr{P}_{Z,K}, \quad Z_K \in \mathcal{M}_{n,K},$$
$$J^{(k,l)} \equiv J^{(k,l)}(Z,K) = J_{k,l}(Z,K) \times J_{l,k}(Z,K).$$

and then find $\hat{K}$ as

$$\hat{K} \in \underset{K}{\operatorname{argmin}}\left\{\sum_{k,l=1}^{K}\left\|A^{(k,l)}(\hat{Z}_K,K) - \Pi_{(1)}\left(\Pi_{\hat{J}_K^{(k,l)}}\left(A^{(k,l)}(\hat{Z}_K,K)\right)\right)\right\|_F^2 + \text{Pen}(n,\hat{J}_K,K)\right\}. \tag{4.9}$$

4.2.3 The Support of the Probability Matrix and the Penalty

Consider solution of optimization problem (4.8) for a fixed value of K. If $\hat{Z}_K \in \mathcal{M}_{n,K}$ is a solution of (4.7), then

$$\hat{J}_K \in \underset{J}{\operatorname{argmin}} \left\{ \sum_{k,l=1}^{K} \left\| A^{(k,l)}(\hat{Z}_K, K) - \Pi_{(1)} \left(\Pi_{J^{(k,l)}} \left(A^{(k,l)}(\hat{Z}_K, K) \right) \right) \right\|_F^2 + \text{Pen}(n, J, K) \right\}$$
$$\text{s.t.} \quad A(\hat{Z}_K, K) = \mathscr{P}^T_{\hat{Z}_K,K} A \mathscr{P}_{\hat{Z}_K,K}, \quad J^{(k,l)} = J_{k,l} \times J_{l,k}, \quad J_{k,l} \equiv J_{k,l}(\hat{Z}_K, K). \tag{4.10}$$

Observe that if the penalty term $\text{Pen}(n, J, K)$ were not present in (4.10) or did not depend on set J, then one would have $\hat{J}_K = \breve{J}_K$ and $\hat{J}_K^{(k,l)} = \breve{J}_K^{(k,l)}$ where, by (4.2), $\breve{J}_K^{(k,l)}$ is the set of indices of nonzero rows and columns in $A^{(k,l)}(\hat{Z}_K, K)$. It is easy to see that

$$\Pi_{\breve{J}^{(k,l)}} \left(A^{(k,l)}(\hat{Z}_K, K) \right) = A^{(k,l)}(\hat{Z}_K, K),\ \Pi_{(1)} \left(\Pi_{\breve{J}^{(k,l)}} \left(A^{(k,l)}(\hat{Z}_K, K) \right) \right) = \Pi_{(1)} \left(A^{(k,l)}(\hat{Z}_K, K) \right).$$

Hence, even if sparsity is not specifically enforced (as it happens in Noroozi et al. (2019*a*) where the penalty depends on n and K only), one still obtains a sparse estimator $\hat{P}$ with the support $\hat{J}_K = \breve{J}_K$.

If the true number of clusters K_* and the true clustering matrix $Z_* \in \mathcal{M}_{n,K_*}$ were available, then the statement below shows that, under certain conditions, with high probability, sets $J_* \equiv J_*(Z_*, K_*)$ and $\breve{J}(Z_*, K_*)$ would coincide.

Lemma 4.2.1. *Let $K_*^2 \leq n$ and the true matrix P_* be such that $(P_*)_{i,j} = 0$ or $(P_*)_{i,j} > \varpi(n, K_*)$. If the community sizes are balanced, i.e., the sizes of the true communities are no less than $\tilde{C}_0 n / K_*$ for some $\tilde{C}_0 \in (0, 1]$, and*

$$\varpi(n, K_*) \geq K_* \left(\sqrt{\ln n} + \sqrt{t} \right) \Big/ \left(\tilde{C}_0 \sqrt{2n} \right),$$

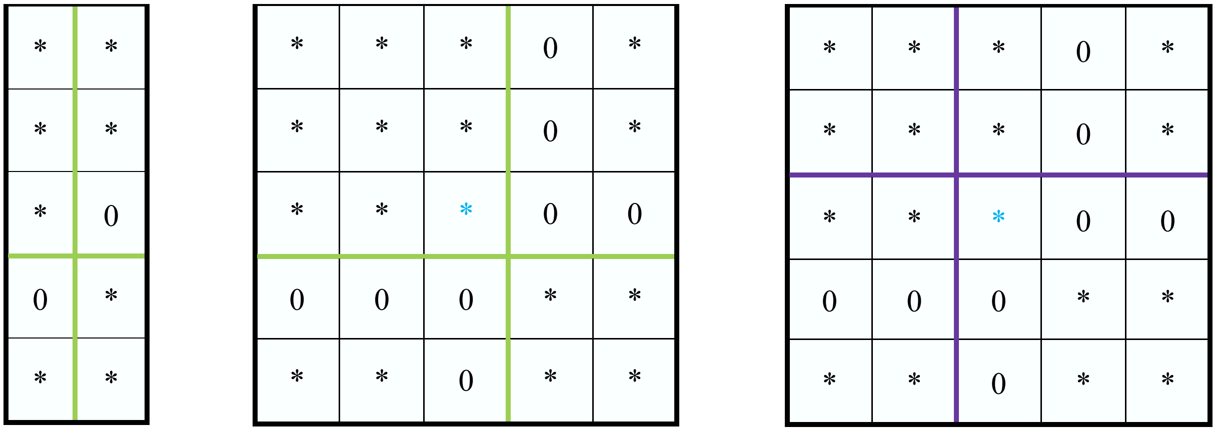

Figure 4.1: Zeros of the probability matrix with $n = 5$ and $K_* = 2$. Star symbols correspond to nonzero elements, the thick lines correspond to clustering assignments. Left panel: matrix Λ with $(J_*)_{1,1} = \{1,2,3\}$, $(J_*)_{2,1} = \{5\}$, $(J_*)_{1,2} = \{1,2\}$ and $(J_*)_{2,2} = \{4,5\}$. Middle panel: matrix $P_*(Z_*, K_*)$ with true clustering, $(\breve{J}_*)^c_{2,1}(Z_*) = \{4\}$ and $(\breve{J}_*)^c_{1,2}(Z_*) = \{3\}$, $\hat{P}_{i,j}(Z_*, K_*) = 0$ for $(i,j) \in \{(1,4),(2,4),(3,4),(3,5),(4,1),(4,2),(4,3),(5,3)\}$, so that, zero entries of the probability matrix are estimated by zeros. Right panel: matrix $P_*(\hat{Z}, K_*)$ with node 3 erroneously placed into community 2. The value of $(P_*)_{3,3}$ is nonzero. If $A_{3,3} = 0$, then $\breve{J}^c_{2,2}(\hat{Z}) = \{3\}$ and $\hat{P}_{i,j}(\hat{Z}, K_*) = 0$ for $(i,j) \in \{(1,4),(2,4),(3,4),(3,5),(4,1),(4,2),(4,3),(5,3)\}$, hence, zero entries of P_* are still estimated by the identical zeros. However, if $A_{3,3} = 1$, then zero elements $(P_*)_{3,4}$, $(P_*)_{3,5}$, $(P_*)_{4,3}$ and $(P_*)_{5,3}$ are estimated by positive values.

then, with probability at least $1 - e^{-t}$, *one has* $J_*(Z_*, K_*) = \breve{J}(Z_*, K_*)$.

Unfortunately, K_* and Z_* are unknown and, hence, $\hat{J}_K(Z,K) = \breve{J}_K(Z,K)$ may not always be the best estimator.

Consider, for example, the situation displayed in Figure 1 where $n = 5$, $K_* = 2$ and, under the true clustering, one has $n_1 = 3$ and $n_2 = 2$. Vectors $\Lambda_{2,1}$ and $\Lambda_{1,2}$ have one zero element each, so that $(J_*)_{1,1} = \{1,2,3\}$, $(J_*)_{2,1} = \{5\}$, $(J_*)_{1,2} = \{1,2\}$ and $(J_*)_{2,2} = \{4,5\}$ (left panel) leading to $(J_*)^{(1,1)} = \{(1,1),(1,2),(1,3),(2,1),(2,2),(2,3),(3,1),(3,2),(3,3)\}$, $(J_*)^{(2,1)} = \{(5,1),(5,2)\}$, $(J_*)^{(1,2)} = \{(1,5),(2,5)\}$ and $(J_*)^{(2,2)} = \{(4,4),(4,5),(5,4),(5.5)\}$ (middle

panel). With the true clustering (middle panel), $(\breve{J}_*)^c_{2,1}(Z_*) = \{4\}$ and $(\breve{J}_*)^c_{1,2}(Z_*) = \{3\}$, so that $\hat{P}_{i,j}(Z_*, K_*) = 0$ for $(i,j) \in \{(1,4),(2,4),(3,4),(3,5),(4,1),(4,2),(4,3),(5,3)\}$. Hence, zero entries of the probability matrix are estimated by zeros. Consider now the situation where the third node has been erroneously placed into community 2 by clustering matrix $\hat{Z}$ (right panel). Then, we still have $(\breve{J}_*)^c_{2,1}(\hat{Z}) = \{4\}$, but $(\breve{J}_*)^c_{1,2}(\hat{Z})$ is an empty set. If $A_{3,3} = 0$, then $\breve{J}^c_{2,2}(\hat{Z}) = \{3\}$ and $\hat{P}_{i,j}(\hat{Z}, K_*) = 0$ for $(i,j) \in \{(1,4),(2,4),(3,4),(3,5),(4,1),$
$(4,2),(4,3),(5,3)\}$, so that the zero entries of P_* are still estimated by the identical zeros. However, if $A_{3,3} = 1$, then zero elements $(P_*)_{3,4}$, $(P_*)_{3,5}$, $(P_*)_{4,3}$ and $(P_*)_{5,3}$ will be estimated by positive values.

For this reason, it is reasonable to introduce a penalty that will lead to trimming the support of $\hat{P}(Z,K)$.

We say that a penalty $Pen(n,J,K)$ is *separable* if for any K and any clustering matrix Z that partitions n nodes into K communities of sizes n_k, $k = 1,\dots,K$, one can write

$$\text{Pen}(n,J,K) = \text{Pen}^{(0)}(n,J,K) + \text{Pen}^{(1)}(n,K) \quad \text{with} \quad \text{Pen}^{(0)}(n,J,K) = \sum_{l=1}^{K}\sum_{k=1}^{K} \mathscr{F}(|J_{k,l}|, n_k), \tag{4.11}$$

where $J_{k,l} \equiv J_{k,l}(Z,K)$. Otherwise, the penalty is *non-separable*.

Lemma 4.2.2. *Let $(\hat{Z}_K, \hat{J}_K)$ be the solution of the optimization problem* (4.8). *If $Pen(n,J,K)$ is separable and function $\mathscr{F}(j,m)$ in* (4.11) *is an increasing function of j for $0 \le j \le m$, then, for any $K < n$ and $k,l = 1,\dots,K$, one has*

$$\hat{J}_{k,l}(\hat{Z}_K, K) \subseteq \breve{J}_{k,l}(\hat{Z}_K, K) \subseteq (\breve{J}_*)_{k,l}(\hat{Z}_K, K), \quad \hat{J}(\hat{Z}_K, K) \subseteq \breve{J}(\hat{Z}_K, K) \subseteq \breve{J}_*(\hat{Z}_K, K). \tag{4.12}$$

4.3 The Errors of Estimation and Clustering

We produce upper bounds on the estimation and clustering errors in this section.

4.3.1 *The penalty*

In what follows, we consider the separable and the non-separable penalties of the form (4.11) with the common $\text{Pen}^{(1)}(n, K)$, i.e.

$$\text{Pen}^{(a)}(n, J, K) = \text{Pen}^{(0,a)}(n, J, K) + \text{Pen}^{(1)}(n, K), \tag{4.13}$$

where a =s for the separable penalty and a = ns for the nonseparable one, and

$$\text{Pen}^{(0,s)}(n, J, K) = \beta_1 \sum_{k,l=1}^{K} |J_{k,l}| \ln(n_k e/|J_{k,l}|) + \beta_2 K \sum_{k=1}^{K} \ln n_k \tag{4.14}$$

$$\text{Pen}^{(0,ns)}(n, J, K) = \beta_1 |J| \ln(nKe/|J|) + 2\beta_2 \ln n \tag{4.15}$$

$$\text{Pen}^{(1)}(n, K) = \beta_2[n \ln K + \ln n]. \tag{4.16}$$

Here, the separable penalty corresponds to $\mathscr{F}(|J_{k,l}|, n_k) = \beta_1 |J_{k,l}| \ln(n_k e/|J_{k,l}|) + \beta_2 \ln n_k$ and the exact expressions for β_1 and β_2 are given in Theorem 3.6.1 below.

In the next two sections, we shall provide upper bounds for the errors of the solution of optimization problem (4.5) with the separable or the non-separable penalty as well as upper bounds for the clustering error in the case of the separable penalty. While the separable penalty has some valuable properties (see Lemma 4.2.2), the non-separable penalty is much easier to interpret. Fortunately, as the statement below shows, under very nonrestrictive conditions, the penalties are within a constant factor of each other.

Lemma 4.3.1. *If $n \geq 8$ and $K \leq \sqrt{n/\ln n}$, then*

$$\textit{Pen}^{(ns)}(n, J, K) < (2 + \beta_1/\beta_2)\, \textit{Pen}^{(s)}(n, J, K) < 2\,(2 + \beta_1/\beta_2)\, \textit{Pen}^{(ns)}(n, J, K). \tag{4.17}$$

4.3.2 The Estimation Errors

Theorem 4.3.1. *Let $(\hat{\Theta}, \hat{Z}, \hat{J}, \hat{K})$ be a solution of optimization problem* (4.5) *with the separable or non-separable penalty defined in* (4.13). *Construct the estimator $\hat{P}$ of P_* of the form*

$$\hat{P} = \mathscr{P}_{\hat{Z},\hat{K}} \hat{\Theta}(\hat{Z}, \hat{J}, \hat{K}) \mathscr{P}^T_{\hat{Z},\hat{K}} \tag{4.18}$$

where $\mathscr{P}_{\hat{Z},\hat{K}}$ is the permutation matrix corresponding to $(\hat{Z}, \hat{K})$. Let positive γ_1, γ_2 be such that $\gamma_1 + \gamma_2 < 1$ and β_1 and β_2 in (4.14)–(4.16) *be given by*

$$\beta_1 = \frac{2(C_1 + C_2)(8 + \gamma_1)}{\gamma_1} + \frac{2}{\gamma_2}, \quad \beta_2 = \frac{2C_2(8 + \gamma_1)}{\gamma_1} + \frac{2}{\gamma_2}, \tag{4.19}$$

where C_1 and C_2 are absolute constants in Lemma 4.4.3. Then, for any $t > 0$, one has

$$\mathbb{P}\left\{ \frac{1}{n^2} \left\| \hat{P} - P_* \right\|_F^2 \leq \frac{\textit{Pen}(n, J_*, K_*)}{n^2\,(1 - \gamma_1 - \gamma_2)} + \frac{\tilde{C} t}{n^2} \right\} \geq 1 - 3e^{-t}, \tag{4.20}$$

and,

$$\frac{1}{n^2}\, \mathbb{E} \left\| \hat{P} - P_* \right\|_F^2 \leq \frac{\textit{Pen}(n, J_*, K_*)}{n^2\,(1 - \gamma_1 - \gamma_2)} + \frac{3\tilde{C}}{n^2} \tag{4.21}$$

where

$$\tilde{C} = 2\,{\gamma_1}^{-1}{\gamma_2}^{-1}(1 - \gamma_1 - 8\gamma_2)^{-1}\,(C_2\gamma_1\gamma_2 + \gamma_1 + 8C_2\gamma_2) \tag{4.22}$$

Observe that, due to Lemma 4.3.1, the separable and non-separable penalties are within a constant factor of each other, so that Theorem 4.3.1 implies that the estimation error is proportional to Pen(n, J_*, K_*) where

$$\text{Pen}(n, J, K) \asymp \text{Pen}^{(ns)}(n, J, K) \asymp n \ln K + |J| \, \ln(nKe/|J|) + \ln n. \tag{4.23}$$

The first term in (4.23) is due to the clustering errors, the second term quantifies the difficulty of finding and estimating $|J|$ nonzero elements among nK elements of matrix $\Lambda \in [0, 1]^{n \times K}$ while the $\ln n \asymp \ln(nK)$ term stands for the difficulty of finding the cardinality of the set $|J|$, and it is always dominated by the first two terms in (4.23).

Since each node has at least one community to which it is connected with a nonzero probability, one has $n \le |J| \le nK$. In the (non-sparse) PABM, $|J| = nK$ and the second term in (4.23) is always asymptotically larger, as $n \to \infty$, than the other two terms. In SPABM, the second term in (4.23) dominates the first term only if $K = 1$ or $|J|/n \to \infty$ as $n \to \infty$. However, if $K > 1$ and $|J| \asymp n$, then both terms are of the equal asymptotic order. If $K \to \infty$ and $|J| \asymp n$ as $n \to \infty$, then SPABM has the error $O(n \ln K)$ which is asymptotically smaller than $O(nK)$ error of PABM.

4.3.3 The Clustering Errors

In order to evaluate the clustering error, we assume that the true number of classes $K = K_*$ is known. Let $Z_* \in \mathcal{M}_{n,K_*}$ be the true clustering matrix. Then $\hat{Z} \equiv \hat{Z}_K$ is a solution of the optimization problem (4.8). Note that if Z_* is the true clustering matrix and Z is any other clustering matrix, then the proportion of misclustered nodes can be evaluated as

$$\text{Err}(Z, Z_*) = (2n)^{-1} \min_{\mathscr{P}_K \in \mathcal{P}_K} \|Z \mathscr{P}_K - Z_*\|_1 = (2n)^{-1} \min_{\mathscr{P}_K \in \mathcal{P}_K} \|Z \mathscr{P}_K - Z_*\|_F^2 \tag{4.24}$$

where $\mathcal{P}_K$ is the set of permutation matrices $\mathscr{P}_K : \{1, 2, \cdots, K\} \longrightarrow \{1, 2, \cdots, K\}$.

Theorem 4.3.2. *Let $K = K_*$ be the true number of clusters and $Z_* \in \mathcal{M}_{n,K_*}$ be the true clustering matrix and n_k be the true number of nodes in cluster $k = 1, \ldots, K$. Denote by $\gamma(Z_*, \rho_n)$ the set of clustering matrices with the proportion of at most ρ_n of the mis-clustered nodes. Let P_* and $J_* = J_*(P_*, Z_*)$ be, respectively, the true probability matrix and the true set J_*. If for some $\gamma_1, \gamma_2 > 0$ such that $\gamma_1 + \gamma_2 < 1$ and some $\tau \in (0,1)$, one has*

$$\max_{\hat{Z} \in \gamma(Z_*, \rho_n)} \left\{ \sum_{k,l=1}^{K} \|P_*^{(k,l)}(\hat{Z})\|_{op}^2 - \frac{2C_1(\beta_1 - C_1 - C_2)}{(C_1 + C_2)\beta_1 \gamma_2} K \sum_{k=1}^{K} \ln(\hat{n}_k) \right\}$$
$$\leq \frac{(1-\tau)(\beta_1 - C_1 - C_2)}{\beta_1} \left[\|P_*\|_F^2 - 2(1+\sqrt{2})^2 \tau^{-1} \left(C_1 |J_*| + C_2 t \right) \right] \tag{4.25}$$
$$- (\beta_1 - C_1 - C_2) \left[\frac{C_2}{C_1 + C_2} (n \ln K + t) + \sum_{k,l=1}^{K} |(J_*)_{k,l}| \ln \left(\frac{n_k\, e}{|(J_*)_{k,l}|} \right) + \frac{\beta_2}{\beta_1} K \sum_{k=1}^{K} \ln(n_k) \right]$$

where β_1 and β_2 are defined in (4.19)*, then with probability at least $1 - 2\exp(-t)$, the proportion of mis-clustered nodes does not exceed ρ_n.*

4.4 Proofs

This section consists of the detailed proofs of the results in the SPABM starting from the proofs of the main results and then the supplementary lemmas.

4.4.1 Proof of Theorem 4.3.1

In what follows, $F_j(n,J,K)$ will stand for $F_j^{(s)}(n,J,K)$ or $F_j^{(ns)}(n,J,K)$, $j=1,2$, where

$$F_1^{(ns)}(n,J,K) = (C_1+C_2)|J|\ln(nKe/|J|) + C_2(3\ln n + n\ln K) \tag{4.26}$$

$$F_1^{(s)}(n,J,K) = (C_1+C_2)\sum_{k,l=1}^{K}|J_{k,l}|\ln(n_k e/|J_{k,l}|) + C_2\left(\ln n + n\ln K + K\sum_{k=1}^{K}\ln n_k\right) \tag{4.27}$$

$$F_2^{(ns)}(n,J,K) = 2\ln n + 2(n+2)\ln K + 2|J|\ln(nKe/|J|) \tag{4.28}$$

$$F_2^{(s)}(n,J,K) = 2\sum_{k,l=1}^{K}|J_{k,l}|\ln(n_k e/|J_{k,l}|) + 2\left(\ln n + n\ln K + K\sum_{k=1}^{K}\ln n_k\right) \tag{4.29}$$

Denote $\Xi = A - P_*$ and recall that, given matrix P_*, entries $\Xi_{i,j} = A_{i,j} - (P_*)_{ij}$ of Ξ are the independent Bernoulli errors for $1 \le i \le j \le n$ and $\Xi_{i,j} = \Xi_{j,i}$.

Let $(\hat{\Theta}, \hat{Z}, \hat{J}, \hat{K})$ be a solution of optimization problem (4.5). We construct the estimator $\hat{P} \equiv \hat{P}(\hat{Z}, \hat{J}, \hat{K})$ of P_* of the form (4.18). Since $A(Z,K) = \mathscr{P}_{Z,K}^T A \mathscr{P}_{Z,K}$, then $A = \mathscr{P}_{Z,K} A(Z,K) \mathscr{P}_{Z,K}^T$, and $\hat{\Theta}(\hat{Z}, \hat{J}, \hat{K})$ is the block matrix of optimal rank one approximations for every block of $\Pi_{\hat{J}}(A(\hat{Z}, \hat{K}))$. Then (4.5) yields

$$\left\|\mathscr{P}_{\hat{Z},\hat{K}}^T A \mathscr{P}_{\hat{Z},\hat{K}} - \hat{\Theta}(\hat{Z},\hat{J},\hat{K})\right\|_F^2 + \text{Pen}(n,\hat{J},\hat{K}) \le \left\|\mathscr{P}_{Z_*}^T A \mathscr{P}_{Z_*}^T - \mathscr{P}_{Z_*}^T P_* \mathscr{P}_{Z_*}^T\right\|_F^2 + \text{Pen}(n,J_*,K_*)$$

Using orthogonality of permutation matrices, obtain

$$\left\|A - \mathscr{P}_{\hat{Z},\hat{K}}\hat{\Theta}(\hat{Z},\hat{J},\hat{K})\mathscr{P}_{\hat{Z},\hat{K}}^T\right\|_F^2 \le \|A-P_*\|_F^2 + \text{Pen}(n,J_*,K_*) - \text{Pen}(n,\hat{J},\hat{K}) \tag{4.30}$$

Hence (4.30) and (4.18) yield

$$\left\|A - \hat{P}\right\|_F^2 \leq \|A - P_*\|_F^2 + \text{Pen}(n, J_*, K_*) - \text{Pen}(n, \hat{J}, \hat{K}) \tag{4.31}$$

Now adding and subtracting P_* in the norm on the left side of (4.31), we rewrite (4.31) as

$$\left\|\hat{P} - P_*\right\|_F^2 \leq \Delta(\hat{Z}, \hat{J}, \hat{K}) + \text{Pen}(n, J_*, K_*) - \text{Pen}(n, \hat{J}, \hat{K}) \tag{4.32}$$

where

$$\Delta(\hat{Z}, \hat{J}, \hat{K}) = 2\text{Tr}\left[(A - P_*)^T(\hat{P}(\hat{Z}, \hat{J}, \hat{K}) - P_*)\right].$$

Again using orthogonality of permutation matrices, we can rewrite

$$\Delta(\hat{Z}, \hat{J}, \hat{K}) = 2\langle \Xi(\hat{Z}, \hat{K}), (\hat{\Theta}(\hat{Z}, \hat{J}, \hat{K}) - P_*(\hat{Z}, \hat{K}))\rangle$$

where $\langle A, B\rangle = \text{Tr}(A^T B)$.

Let

Then, in the block form, $\Delta(\hat{Z}, \hat{J}, \hat{K})$ appears as

$$\Delta(\hat{Z}, \hat{J}, \hat{K}) = \sum_{k,l=1}^{\hat{K}} \Delta^{(k,l)}(\hat{Z}, \hat{J}, \hat{K}) \tag{4.33}$$

with

$$\Delta^{(k,l)}(\hat{Z}, \hat{J}, \hat{K}) = 2\left\langle \Xi^{(k,l)}(\hat{Z}, \hat{K}), \Pi_{\hat{u},\hat{v}}\left(\Pi_{\hat{J}(k,l)}\left(A^{(k,l)}(\hat{Z}, \hat{K})\right)\right) - P_*^{(k,l)}(\hat{Z}, \hat{K})\right\rangle.$$

Here, $\hat{u} \equiv \hat{u}^{(k,l)}(\hat{Z}, \hat{J}, \hat{K})$ and $\hat{v} \equiv \hat{v}^{(k,l)}(\hat{Z}, \hat{J}, \hat{K})$ are the singular vectors of $\Pi_{\hat{J}(k,l)}(A^{(k,l)}(\hat{Z}, \hat{K}))$

corresponding to the largest singular values of $\Pi_{\hat{J}(k,l)}(A^{(k,l)}(\hat{Z},\hat{K}))$, and $\Pi_{\hat{u},\hat{v}}$ is defined in (3.5).

Recall that

$$\Pi_{\hat{u},\hat{v}}(\Pi_{\hat{J}(k,l)}(A^{(k,l)}(\hat{Z},\hat{K}))) = \Pi_{\hat{u},\hat{v}}(\Pi_{\hat{J}(k,l)}(P_*^{(k,l)}(\hat{Z},\hat{K})) + \Pi_{\hat{J}(k,l)}(\Xi^{(k,l)}(\hat{Z},\hat{K})))$$

Hence, $\Delta^{(k,l)}(\hat{Z},\hat{J},\hat{K})$ can be partitioned into the sums of three components

$$\Delta^{(k,l)}(\hat{Z},\hat{J},\hat{K}) = \Delta_1^{(k,l)}(\hat{Z},\hat{J},\hat{K}) + \Delta_2^{(k,l)}(\hat{Z},\hat{J},\hat{K}) + \Delta_3^{(k,l)}(\hat{Z},\hat{J},\hat{K}), \quad k,l = 1,2,\cdots,K, \tag{4.34}$$

where

$$\Delta_1^{(k,l)}(\hat{Z},\hat{J},\hat{K}) = 2\left\langle \Xi^{(k,l)}(\hat{Z},\hat{K}), \Pi_{\hat{u},\hat{v}}(\Pi_{\hat{J}(k,l)}(\Xi^{(k,l)}(\hat{Z},\hat{K}))) \right\rangle \tag{4.35}$$

$$\Delta_2^{(k,l)}(\hat{Z},\hat{J},\hat{K}) = 2\left\langle \Xi^{(k,l)}(\hat{Z},\hat{K}), \Pi_{\tilde{u},\tilde{v}}(\Pi_{\hat{J}(k,l)}({P_*}^{(k,l)}(\hat{Z},\hat{K}))) - P_*^{(k,l)}(\hat{Z},\hat{K}) \right\rangle \tag{4.36}$$

$$\Delta_3^{(k,l)}(\hat{Z},\hat{J},\hat{K}) = 2\left\langle \Xi^{(k,l)}(\hat{Z},\hat{K}), \Pi_{\hat{u},\hat{v}}(\Pi_{\hat{J}(k,l)}({P_*}^{(k,l)}(\hat{Z},\hat{K}))) - \Pi_{\tilde{u},\tilde{v}}(\Pi_{\hat{J}(k,l)}(P_*^{(k,l)}(\hat{Z},\hat{K}))) \right\rangle. \tag{4.37}$$

Here $\tilde{u} = \tilde{u}^{(k,l)}(\hat{Z},\hat{J},\hat{K})$ and $\tilde{v} = \tilde{v}^{(k,l)}(\hat{Z},\hat{J},\hat{K})$ are the singular vectors of $\Pi_{\hat{J}(k,l)}({P_*}^{(k,l)}(\hat{Z},\hat{K}))$ corresponding to the largest singular values of $\Pi_{\hat{J}(k,l)}({P_*}^{(k,l)}(\hat{Z},\hat{K}))$ and $\Pi_{\tilde{u},\tilde{v}}(\Pi_{\hat{J}(k,l)}({P_*}^{(k,l)}(\hat{Z},\hat{K})))$ is defined in (3.5). With some abuse of notations, for any matrix B and any vectors u, v, let $\Pi_{u,v}\left(\Pi_{\hat{J}}(B(\hat{Z},\hat{K}))\right)$ be the matrix with blocks $\Pi_{u,v}\left(\Pi_{\hat{J}(k,l)}(B^{(k,l)}(\hat{Z},\hat{K}))\right)$, $k,l = 1,2,\cdots,\hat{K}$. Then, it follows from (4.34)–(4.37) that

$$\Delta(\hat{Z},\hat{J},\hat{K}) = \Delta_1(\hat{Z},\hat{J},\hat{K}) + \Delta_2(\hat{Z},\hat{J},\hat{K}) + \Delta_3(\hat{Z},\hat{J},\hat{K}) \tag{4.38}$$

where

$$\Delta_1(\hat{Z},\hat{J},\hat{K}) = 2\left\langle \Xi(\hat{Z},\hat{K}), \Pi_{\hat{u},\hat{v}}\left(\Pi_{\hat{J}}(\Xi(\hat{Z},\hat{K}))\right)\right\rangle \tag{4.39}$$

$$\Delta_2(\hat{Z},\hat{J},\hat{K}) = 2\left\langle \Xi(\hat{Z},\hat{K}), \Pi_{\tilde{u},\tilde{v}}\left(\Pi_{\hat{J}}(P_*(\hat{Z},\hat{K}))\right) - P_*(\hat{Z},\hat{K})\right\rangle \tag{4.40}$$

$$\Delta_3(\hat{Z},\hat{J},\hat{K}) = 2\left\langle \Xi(\hat{Z},\hat{K}), \Pi_{\hat{u},\hat{v}}\left(\Pi_{\hat{J}}(P_*(\hat{Z},\hat{K}))\right) - \Pi_{\tilde{u},\tilde{v}}\left(\Pi_{\hat{J}}(P_*(\hat{Z},\hat{K}))\right)\right\rangle \tag{4.41}$$

Now, we need to derive an upper bound for each component in (4.38).

Observe that

$$|\Delta_1^{(k,l)}(\hat{Z},\hat{J},\hat{K})| = 2\left\|\Pi_{\hat{u},\hat{v}}(\Pi_{\hat{J}(k,l)}(\Xi^{(k,l)}(\hat{Z},\hat{K})))\right\|_F^2 \le 2\left\|\Pi_{\hat{J}(k,l)}(\Xi^{(k,l)}(\hat{Z},\hat{K}))\right\|_F^2$$

Fix $t > 0$ and let Ω_1 be the set such that $\|\Pi_{\hat{J}}\left(\Xi(\hat{Z},\hat{K})\right)\|_{op}^2 \le F_1(n,\hat{J},\hat{K}) + C_2 t$. According to Lemma 4.4.4,

$$\mathbb{P}(\Omega_1) \ge 1 - \exp(-t), \tag{4.42}$$

and, for $\omega \in \Omega_1$, one has

$$|\Delta_1(\hat{Z},\hat{J},\hat{K})| \le 2\sum_{k,l=1}^{\hat{K}} \|\Pi_{\hat{J}(k,l)}(\Xi^{(k,l)}(\hat{Z},\hat{K}))\|_{op}^2 \le 2\,F_1(n,\hat{J},\hat{K}) + 2\,C_2 t \tag{4.43}$$

where $F_1(n,J,K)$ is defined by either (4.26) or (4.27) and C_2 is given in Lemma 4.4.3.

Now, derive an upper bound for $\Delta_2(\hat{Z},\hat{J},\hat{K})$ given by (4.40). Note that

$$|\Delta_2(\hat{Z},\hat{J},\hat{K})| = 2\|\Pi_{\tilde{u},\tilde{v}}\left(\Pi_{\hat{J}}(P_*(\hat{Z},\hat{K}))\right) - P_*(\hat{Z},\hat{K})\|_F\; |\langle \Xi(\hat{Z},\hat{K}), H_{\tilde{u},\tilde{v}}(\hat{Z},\hat{J},\hat{K})\rangle|,$$

where

$$H_{\tilde{u},\tilde{v}}(\hat{Z},\hat{J},\hat{K}) = \frac{\Pi_{\tilde{u},\tilde{v}}\left(\Pi_{\hat{J}}(P_*(\hat{Z},\hat{K}))\right) - P_*(\hat{Z},\hat{K})}{\|\Pi_{\tilde{u},\tilde{v}}\left(\Pi_{\hat{J}}(P_*(\hat{Z},\hat{K}))\right) - P_*(\hat{Z},\hat{K})\|_F}$$

Since for any a, b and $\alpha_1 > 0$, one has $2ab \le \alpha_1 a^2 + b^2/\alpha_1$, obtain

$$|\Delta_2(\hat{Z},\hat{J},\hat{K})| \le \alpha_1 \|\Pi_{\tilde{u},\tilde{v}}\left(\Pi_{\hat{J}}(P_*(\hat{Z},\hat{K}))\right) - P_*(\hat{Z},\hat{K})\|_F^2 + \frac{1}{\alpha_1}\,|\langle \Xi(\hat{Z},\hat{K}), H_{\tilde{u},\tilde{v}}(\hat{Z},\hat{J},\hat{K})\rangle|^2 \tag{4.44}$$

Observe that if K, J and $Z \in \mathcal{M}_{n,K}$ are fixed, then $H_{\tilde{u},\tilde{v}}(Z,J,K)$ is fixed and, for any K, J and Z, one has $\|H_{\tilde{u},\tilde{v}}(Z,J,K)\|_F = 1$. Note also that, for fixed K, J and Z, matrix $\Xi(Z,K) \in [0,1]^{n\times n}$ contains independent Bernoulli errors. It is well known that if ξ is a vector of independent Bernoulli errors and h is any fixed vector, then, for any $x > 0$, by Hoeffding's inequality $\mathbb{P}(|\xi^T h|^2 > x) \le 2\exp(-x/2)$. Since $\langle \Xi(Z,K), H_{\tilde{u},\tilde{v}}(Z,J,K)\rangle = [\text{vec}(\Xi(Z,K))]^T \text{vec}(H_{\tilde{u},\tilde{v}}(Z,J,K))$, obtain for any fixed K, J and Z

$$\mathbb{P}\left(|\langle \Xi(Z,K), H_{\tilde{u},\tilde{v}}(Z,J,K)\rangle|^2 - x > 0\right) \le 2\exp(-x/2).$$

Hence, application of the union bound yields

$$\begin{aligned} &\mathbb{P}\left(|\langle \Xi(\hat{Z},\hat{K}), H_{\tilde{u},\tilde{v}}(\hat{Z},\hat{J},\hat{K})\rangle|^2 - F_2(n,\hat{J},\hat{K}) > 2t\right) \\ &\le \mathbb{P}\left(\max_{1\le K\le n}\max_{J}\max_{Z\in\mathcal{M}_{n,k}}[|\langle \Xi(Z,K), H_{\tilde{u},\tilde{v}}(Z,J,K)\rangle|^2 - F_2(n,J,K)] > 2\,t\right) \le 2\,\exp(-t), \end{aligned} \tag{4.45}$$

where $F_2(n,\hat{J},\hat{K})$ is defined by (4.28) or (4.29). Using Lemma 4.4.2, obtain that

$$\|\Pi_{\tilde{u},\tilde{v}}\left(\Pi_{\hat{J}}(P_*(\hat{Z},\hat{K}))\right) - P_*(\hat{Z},\hat{K})\|_F^2 \le \|\Pi_{\hat{u},\hat{v}}\left(\Pi_{\hat{J}}(P_*(\hat{Z},\hat{K}))\right) - P_*(\hat{Z},\hat{K})\|_F^2 \le \|\hat{P}(\hat{Z},\hat{J},\hat{K}) - P_*\|_F^2.$$

Denote the set on which (4.45) holds by Ω_2^c, so that

$$\mathbb{P}(\Omega_2) \geq 1 - 2\exp(-t). \tag{4.46}$$

Then inequalities (4.44) and (4.45) imply that, for any $\alpha_1 > 0$ and any $\omega \in \Omega_2$, one has

$$|\Delta_2(\hat{Z}, \hat{J}, \hat{K})| \leq \alpha_1 \|\hat{P}(\hat{Z}, \hat{J}, \hat{K}) - P_*\|_F^2 + 1/\alpha_1\, F_2(n, \hat{J}, \hat{K}) + 2\, t/\alpha_1. \tag{4.47}$$

Now consider $\Delta_3(\hat{Z}, \hat{J}, \hat{K})$ defined in (4.41) with components (4.37). Note that matrices $X_{k,l} = \Pi_{\hat{u},\hat{v}}(\Pi_{\hat{J}(k,l)}({P_*}^{(k,l)}(\hat{Z}, \hat{K}))) - \Pi_{\tilde{u},\tilde{v}}(\Pi_{\hat{J}(k,l)}(P_*^{(k,l)}(\hat{Z}, \hat{K})))$ have ranks at most two. Use the fact that (see, e.g., Giraud (2014), page 123)

$$\langle A, B\rangle \leq \|A\|_{(2,r)} \|B\|_{(2,r)} \leq r\, \|A\|_{op} \|B\|_F, \quad r = \min\{\text{rank}(A), \text{rank}(B)\}, \tag{4.48}$$

where, for any matrix X, $\|X\|_{(2,q)}$ is the Ky-Fan $(2, q)$ norm such that $\|X\|_{(2,q)}^2 \leq \text{rank}(X)\, \|X\|_{op}^2$. Applying inequality (4.48) with $r = 2$ to $X_{k,l}$ above, derive that

$$|\Delta_3^{(k,l)}(\hat{Z}, \hat{J}, \hat{K})| \leq 4\, \|\Pi_{\hat{J}(k,l)}(\Xi^{(k,l)}(\hat{Z}, \hat{K}))\|_{op} \, \left\|\Pi_{\hat{u},\hat{v}}(\Pi_{\hat{J}(k,l)}({P_*}^{(k,l)}(\hat{Z}, \hat{K}))) - \Pi_{\tilde{u},\tilde{v}}(\Pi_{\hat{J}(k,l)}(P_*^{(k,l)}(\hat{Z}, \hat{K})))\right\|_F$$

Then, for any $\alpha_2 > 0$, obtain

$$|\Delta_3(\hat{Z}, \hat{J}, \hat{K})| = \sum_{k,l=1}^{\hat{K}} |\Delta_3^{(k,l)}(\hat{Z}, \hat{J}, \hat{K})| \leq \frac{2}{\alpha_2} \sum_{k,l=1}^{\hat{K}} \|\Xi^{(k,l)}(\hat{Z}, \hat{K})\|_{op}^2 \tag{4.49}$$

$$+ 2\, \alpha_2 \sum_{k,l=1}^{\hat{K}} \|\Pi_{\hat{u},\hat{v}}(\Pi_{\hat{J}(k,l)}({P_*}^{(k,l)}(\hat{Z}, \hat{K}))) - \Pi_{\tilde{u},\tilde{v}}(\Pi_{\hat{J}(k,l)}(P_*^{(k,l)}(\hat{Z}, \hat{K})))\|_F^2$$

Note that, by Lemma 4.4.2,

$$\| \Pi_{\hat{u},\hat{v}}(\Pi_{\hat{J}(k,l)}(P_*{}^{(k,l)}(\hat{Z},\hat{K}))) - \Pi_{\tilde{u},\tilde{v}}(\Pi_{\hat{J}(k,l)}(P_*^{(k,l)}(\hat{Z},\hat{K})))\|_F^2 \le$$
$$2\,\|\Pi_{\hat{u},\hat{v}}(\Pi_{\hat{J}(k,l)}(P_*{}^{(k,l)}(\hat{Z},\hat{K}))) - P_*^{(k,l)}(\hat{Z},\hat{K})\|_F^2 + 2\,\|\Pi_{\tilde{u},\tilde{v}}(\Pi_{\hat{J}(k,l)}(P_*^{(k,l)}(\hat{Z},\hat{K}))) - P_*^{(k,l)}(\hat{Z},\hat{K})\|_F^2 \le$$
$$4\|\Pi_{\hat{u},\hat{v}}\left(\Pi_{\hat{J}(k,l)}(A^{(k,l)}(\hat{Z},\hat{K}))\right) - P_*^{(k,l)}(\hat{Z},\hat{K})\|_F^2 = 4\|\hat{P} - P_*\|_F^2$$

Combining the last inequality with (4.43) and (4.49), obtain that for any $\alpha_2 > 0$, $t > 0$ and $\omega \in \Omega_1$, one has

$$|\Delta_3(\hat{Z},\hat{J},\hat{K})| \le 8\alpha_2\|\hat{P} - P_*\|_F^2 + 2/\alpha_2 F_1(n,\hat{J},\hat{K}) + 2\,C_2\,t/\alpha_2. \tag{4.50}$$

Let $\Omega = \Omega_1 \cap \Omega_2$. Then, (4.42) and (4.46) imply that $\mathbb{P}(\Omega) \ge 1 - 3\exp(-t)$ and, for $\omega \in \Omega$, inequalities (4.43), (4.47) and (4.50) simultaneously hold. Hence, (4.38) implies that, for any $\omega \in \Omega$,

$$|\Delta(\hat{Z},\hat{J},\hat{K})| \le (2{+}2/\alpha_2)F_1(n,\hat{J},\hat{K}){+}1/\alpha_1\,F_2(n,\hat{J},\hat{K}){+}(\alpha_1{+}8\alpha_2)\|\hat{P}{-}P_*\|_F^2{+}2(C_2{+}1/\alpha_1{+}C_2/\alpha_2)\,t.$$

Combination of the last inequality and (4.32) yields that, for $\alpha_1 + 8\alpha_2 < 1$ and any $\omega \in \Omega$,

$$\begin{aligned}(1-\alpha_1-8\alpha_2)\left\|\hat{P} - P_*\right\|_F^2 \le {} & (2+2/\alpha_2)\,F_1(n,\hat{J},\hat{K}) + 1/\alpha_1\,F_2(n,\hat{J},\hat{K}) \\ & + \ \text{Pen}(n,J_*,K_*) - \text{Pen}(n,\hat{J},\hat{K}) + 2(C_2 + 1/\alpha_1 + C_2/\alpha_2)\,t.\end{aligned} \tag{4.51}$$

Set $\gamma_1 = 8\alpha_2$ and $\gamma_2 = \alpha_1$ and $\text{Pen}(n,\hat{J},\hat{K}) = (2{+}16/\gamma_1)F_1(n,\hat{J},\hat{K}){+}1/\gamma_2 F_2(n,\hat{J},\hat{K})$. Obtain the penalty as defined in (4.13)–(4.16), with the expressions for β_1 and β_2 given in (4.19). Dividing both sides of (4.51) by $(1-\gamma_1-\gamma_2)$, obtain that

$$\mathbb{P}\left\{\|\hat{P} - P_*\|_F^2 \le (1-\gamma_1-\gamma_2)^{-1}\,\text{Pen}(n,J_*,K_*) + \tilde{C}\,t\right\} \ge 1 - 3e^{-t} \tag{4.52}$$

where $\tilde{C}$ is defined in (4.22).

In order to obtain the upper bound (4.21) note that for $\xi = \|\hat{P} - P_*\|_F^2 - (1-\gamma_1-\gamma_2)^{-1}\,\text{Pen}(n, K_*)$, one has $\mathbb{E}\|\hat{P} - P_*\|_F^2 = (1-\gamma_1-\gamma_2)^{-1}\,\text{Pen}(n, K_*) + \mathbb{E}\xi$, where

$$\mathbb{E}\xi \leq \int_0^\infty \mathbb{P}(\xi > z)dz = \tilde{C}\int_0^\infty \mathbb{P}(\xi > \tilde{C}t)dt \leq \tilde{C}\int_0^\infty 3\,e^{-t}\,dt = 3\tilde{C},$$

which yields (4.21).

4.4.2 *Proof of Theorem 4.3.2.*

Let K be fixed, and known so that $K = K_*$ and, hence, $A(\hat{Z}, K) \equiv A(\hat{Z})$ and so on. Let Z_* be the true clustering matrix and J_* be the set of indices such that $P_{i,j}(Z_*, K_*) = 0$ if $(i,j) \notin J_*$. It follows from (4.8) that

$$\begin{aligned}
&\sum_{k,l=1}^{K} \left\| A^{(k,l)}(\hat{Z}) - \Pi_{(1)}(\Pi_{\hat{J}^{(k,l)}}(A^{(k,l)}(\hat{Z}))) \right\|_F^2 + \text{Pen}(n, \hat{J}, K) \\
&\leq \sum_{k,l=1}^{K} \left\| A^{(k,l)}(Z_*) - \Pi_{(1)}(\Pi_{J_*^{(k,l)}}(A^{(k,l)}(Z_*))) \right\|_F^2 + \text{Pen}(n, J_*, K)
\end{aligned}$$

Since for any $Z \in M_{n,K}$ and any J, one has

$$\sum_{k,l=1}^{K} \left\| A^{(k,l)}(Z) \right\|_F^2 = \|A\|_F^2\,, \quad \left\langle A^{(k,l)}(Z), \Pi_{(1)}(\Pi_{J^{(k,l)}}(A^{(k,l)}(Z))) \right\rangle = \|\Pi_{(1)}(\Pi_{J^{(k,l)}}(A^{(k,l)}(Z)))\|_F^2$$

and $\text{Pen}^{(1)}(n, K)$ does not depend on sparsity, obtain

$$\sum_{k,l=1}^{K} \left\| \Pi_{(1)}\left(\Pi_{\hat{J}^{(k,l)}}\left(A^{(k,l)}(\hat{Z})\right)\right) \right\|_F^2 \geq \sum_{k,l=1}^{K} \left\| \Pi_{(1)}\left(\Pi_{J_*^{(k,l)}}\left(A^{(k,l)}(Z_*)\right)\right) \right\|_F^2 \tag{4.53}$$
$$+ \text{Pen}^{(0,s)}(n, \hat{J}, K) - \text{Pen}^{(0,s)}(n, J_*, K).$$

Recall that $P_*^{(k,l)}(Z_*)$ are rank one matrices, while for $Z \neq Z_*$, some $P_*^{(k,l)}(Z)$ may have ranks higher than one. Note that for any $Z \in M_{n,K}$ and any $J^{(k,l)}$

$$\|\Pi_{(1)}(\Pi_{J^{(k,l)}}(A^{(k,l)}(Z)))\|_F \geq \|P_*^{(k,l)}(Z)\|_F - \|\Pi_{(1)}(\Pi_{J^{(k,l)}}(A^{(k,l)}(Z))) - P_*^{(k,l)}(Z)\|_F. \quad (4.54)$$

Denote, as before, $\Xi^{(k,l)}(Z) = A^{(k,l)}(Z) - P_*^{(k,l)}(Z)$. Applying Proposition 6.2 of Giraud (2015) with $\theta = \sqrt{2}$ and $Z = Z_*$ and recalling that matrices $P_*^{(k,l)}(Z_*)$ are of rank one, derive

$$\|\Pi_{(1)}[\Pi_{J^{(k,l)}}(A^{(k,l)}(Z_*))] - P_*^{(k,l)}(Z_*)\|_F^2 \leq 2(1+\sqrt{2})^2 \|\Pi_{J^{(k,l)}}(\Xi^{(k,l)}(Z_*))\|_{op}^2$$

Note that, for $(i,j) \notin J_*^{(k,l)}$, one has $\Xi_{i,j}^{(k,l)}(Z_*) = 0$, so, for any set $J^{(k,l)}$, the matrix $\Pi_{J^{(k,l)}}(\Xi^{(k,l)}(Z_*))$ has $(J_*)_{k,l} \cap J_{k,l}$ nonzero rows and $(J_*)_{l,k} \cap J_{l,k}$ nonzero columns. Therefore, for any $t > 0$, by Lemma 4.4.3

$$\mathbb{P}\left\{ \sum_{k,l=1}^{K} \|\Big(\Pi_{J^{(k,l)}}(\Xi^{(k,l)}(Z_*))\Big)\|_{op}^2 \leq C_1 |J_* \cap J| + C_2\, t \right\} \geq 1 - \exp(-t). \quad (4.55)$$

Observe that, by (4.54), for any $\tau \in (0,1)$, one has

$$\begin{aligned}
&\|\Pi_{(1)}[\Pi_{J_*^{(k,l)}}(A^{(k,l)}(Z_*))]\|_F^2 = \|\Pi_{(1)}(\Pi_{J_*^{(k,l)}}(A^{(k,l)}(Z_*))) - P_*^{(k,l)}(Z_*) + P_*^{(k,l)}(Z_*)\|_F^2 \\
&\geq (1-\tau)\|P_*^{(k,l)}(Z_*)\|_F^2 + (1 - 1/\tau)\|\Pi_{(1)}[\Pi_{J_*^{(k,l)}}(A^{(k,l)}(Z_*))] - P_*^{(k,l)}(Z_*)\|_F^2 \qquad (4.56) \\
&\geq (1-\tau)\|P_*^{(k,l)}(Z_*)\|_F^2 + 2(1+\sqrt{2})^2(1-1/\tau)\,\|\Big(\Pi_{J^{(k,l)}}(\Xi^{(k,l)}(Z_*))\Big)\|_{op}^2.
\end{aligned}$$

Hence, it follows from (4.55) and (4.56), that, for any $\tau \in (0,1)$, any $t > 0$ and $C(\tau) = 2(1+\sqrt{2})^2(1-1/\tau)$

$$\mathbb{P}\left\{ \sum_{k,l=1}^{K} \|\Pi_{(1)}[\Pi_{J_*{}^{(k,l)}}(A^{(k,l)}(Z_*))]\|_F^2 \geq (1-\tau)\|P_*\|_F^2 + C(\tau)\,[C_1|J_*| + C_2\, t] \right\} \geq 1 - e^{-t}. \quad (4.57)$$

On the other hand, for any $\tau_0 \in (0,1)$, derive

$$\|\Pi_{(1)}[\Pi_{\hat{J}^{(k,l)}}(A^{(k,l)}(\hat{Z}))]\|_F^2 \le (1+\tau_0)\|\Pi_{\hat{J}^{(k,l)}} P_*^{(k,l)}(\hat{Z})\|_{op}^2 + (1+1/\tau_0)\|\Pi_{\hat{J}^{(k,l)}} \Xi^{(k,l)}(\hat{Z})\|_{op}^2.$$

Taking a union bound similarly to Lemma 4.4.4 and recalling that K is fixed, obtain for any $t>0$

$$\mathbb{P}\left\{\sum_{k,l=1}^{K} \|\Pi_{\hat{J}^{(k,l)}}(\Xi^{(k,l)}(\hat{Z}))\|_{op}^2 \le [F_1^{(s)}(n,\hat{J},K) - C_2 \ln n] + C_2 t\right\} \ge 1 - e^{-t}$$

where $F_1^{(s)}(n,J,K)$ is defined in (4.27). Therefore, for any $\tau_0 \in (0,1)$ and any $t>0$, derive

$$\mathbb{P}\left\{\sum_{k,l=1}^{K} \|\Pi_{(1)}[\Pi_{\hat{J}^{(k,l)}}(A^{(k,l)}(\hat{Z}))]\|_F^2 \le (1+\tau_0)\sum_{k,l=1}^{K} \|\Pi_{\hat{J}^{(k,l)}} P_*^{(k,l)}(\hat{Z})\|_{op}^2 \right. \tag{4.58}$$
$$\left. + (1+1/\tau_0)\left[(C_1+C_2)\sum_{k,l=1}^{K} |\hat{J}_{k,l}| \ln(\hat{n}_k e/|\hat{J}_{k,l}|) + C_2 n \ln K + C_2 K \sum_{k=1}^{K} \ln(\hat{n}_k) + C_2 t\right]\right\} \ge 1 - e^{-t},$$

where $\hat{n}_k$ is the estimated number of elements in cluster k under clustering matrix $\hat{Z}$. Combining (4.53), (4.57) and (4.58), and plugging expressions for $\text{Pen}^{(0,s)}(n,\hat{J},K)$ and $\text{Pen}^{(0,s)}(n,J_*,K)$, derive that, for any $\tau,\tau_0 \in (0,1)$ and any $t>0$ one has with probability at least $1-2e^{-t}$

$$(1+\tau_0)\sum_{k,l=1}^{K} \|\Pi_{\hat{J}^{(k,l)}} P_*^{(k,l)}(\hat{Z})\|_{op}^2 \ge (1-\tau)\|P_*\|_F^2 + C(\tau)\,[C_1|J_*| + C_2\, t]$$
$$- (1+1/\tau_0)\left[(C_1+C_2)\sum_{k,l=1}^{K} |\hat{J}_{k,l}| \ln(\hat{n}_k e/|\hat{J}_{k,l}|) + C_2 n \ln K + C_2 K \sum_{k=1}^{K} \ln(\hat{n}_k) + C_2 t\right]$$
$$+ \beta_1 \sum_{k,l=1}^{K} |\hat{J}_{k,l}| \ln(\hat{n}_k e/|\hat{J}_{k,l}|) + \beta_2 K \sum_{k=1}^{K} \ln(\hat{n}_k) - \beta_1 \sum_{k,l=1}^{K} |(J_*)_{k,l}| \ln(n_k e/|(J_*)_{k,l}|) - \beta_2 K \sum_{k=1}^{K} \ln(n_k).$$

Recall that, by Lemma 4.2.2, $\hat{J}^{k,l}(\hat{Z}_K, K) \subseteq (\breve{J}_*)^{k,l}(\hat{Z}_K, K)$ for any (k,l), so that

$$\|\Pi_{\hat{J}^{(k,l)}} P_*^{(k,l)}(\hat{Z})\|_{op}^2 \le \|\Pi_{(\breve{J}_*)^{(k,l)}} P_*^{(k,l)}(\hat{Z})\|_{op}^2 = \|P_*^{(k,l)}(\hat{Z})\|_{op}^2.$$

Then, combining the terms, for any $\tau, \tau_0 \in (0,1)$ and any $t > 0$, with probability at least $1 - 2e^{-t}$, arrive at

$$(1+\tau_0) \sum_{k,l=1}^{K} \|\Pi_{\hat{J}^{(k,l)}} P_*^{(k,l)}(\hat{Z})\|_{op}^2 \ge (1-\tau)\|P_*\|_F^2 + C(\tau)\,[C_1|J_*| + C_2\,t] - (1+1/\tau_0)[C_2 n \ln K + C_2 t]$$
$$- \tilde{\beta}_1 \sum_{k,l=1}^{K} |\hat{J}_{k,l}| \ln(\hat{n}_k e/|\hat{J}_{k,l}|) - \tilde{\beta}_2\, K \sum_{k=1}^{K} \ln(\hat{n}_k) - \beta_1 \sum_{k,l=1}^{K} |(J_*)_{k,l}| \ln(n_k e/|(J_*)_{k,l}|) - \beta_2\, K \sum_{k=1}^{K} \ln(n_k),$$

where $\tilde{\beta}_1 = (1+1/\tau_0)(C_1 + C_2) - \beta_1$ and $\tilde{\beta}_2 = (1+1/\tau_0)C_2 - \beta_2$. Choose τ_0 such that $\tilde{\beta}_1 = 0$, then

$$\tilde{\beta}_2 = -\frac{2C_1}{\gamma_2(C_1 + C_2)}, \quad \tau_0 = \frac{C_1 + C_2}{\beta_1 - C_1 - C_2},$$

and recall that $C(\tau) = 2(1+\sqrt{2})^2(1 - 1/\tau)$. Obtain that, for any $\tau, \tau_0 \in (0,1)$ and any $t > 0$, with probability at least $1 - 2e^{-t}$, one has

$$\sum_{k,l=1}^{K} \|P_*^{(k,l)}(\hat{Z})\|_{op}^2 - \frac{2C_1(\beta_1 - C_1 - C_2)}{(C_1+C_2)\beta_1\gamma_2}\, K \sum_{k=1}^{K} \ln(\hat{n}_k)$$
$$\ge \frac{(1-\tau)(\beta_1 - C_1 - C_2)}{\beta_1} \left[\|P_*\|_F^2 - 2(1+\sqrt{2})^2\tau^{-1}\,(C_1|J_*| + C_2 t)\right]$$
$$-\ (\beta_1 - C_1 - C_2)\left[\frac{C_2}{C_1 + C_2}\,(n \ln K + t) + \sum_{k,l=1}^{K} |(J_*)_{k,l}| \ln\left(\frac{n_k\, e}{|(J_*)_{k,l}|}\right) + \frac{\beta_2}{\beta_1} K \sum_{k=1}^{K} \ln(n_k)\right],$$

and the proof is completed by the contradiction argument.

Proof of Lemma 4.2.1. Note that index j is incorrectly identified if $j \in J^*_{l,k} \cap (\breve{J}_{l,k})^c$ or $j \in \breve{J}_{l,k} \cap (J^*_{l,k})^c$. Since Bernoulli variable with zero mean is always equal to zero, the second case is impossible. Observe that for any (k,l), one has $P_*^{(k,l)} \equiv P_*^{(k,l)}(Z_*, K_*)$ and

$$\sum_{i=1}^{n_k} (P_*)_{ij}^{(k,l)} \geq n_k \varpi(n,K) \geq \tilde{C}_0 n K^{-1}\, \varpi(n,K) \text{ if } j \in J^*_{l,k}, \quad \sum_{i=1}^{n_k} (P_*)_{ij}^{(k,l)} = 0 \text{ if } j \in (J^*_{l,k})^c$$

Therefore, for any (k,l) and $j \in J^*_{l,k}$, by Hoeffding inequality,

$$\mathbb{P}(j \in (\breve{J}_{l,k})^c) = \mathbb{P}\left(\sum_{i=1}^{n_k} A_{ij}^{(k,l)}(Z_*, K_*) = 0 \right) = \mathbb{P}\left(\sum_{i=1}^{n_k} \left[A_{ij}^{(k,l)}(Z_*, K_*) - (P_*)_{ij}^{(k,l)} \right] = - \sum_{i=1}^{n_k} (P_*)_{ij}^{(k,l)} \right) \leq$$
$$\mathbb{P}\left(\sum_{i=1}^{n_k} \left[A_{ij}^{(k,l)}(Z_*, K_*) - (P_*)_{ij}^{(k,l)} \right] \leq -\tilde{C}_0 n K_*^{-1}\, \varpi(n, K_*) \right) \leq \exp\left\{ -2 \tilde{C}_0^{\,2} n K_*^{-2}\, \varpi^2(n, K_*) \right\}.$$

Hence, applying the lower bound for $\varpi^2(n, K_*)$ and the union bound, obtain

$$\mathbb{P}(J_*(Z_*, K_*) \neq \breve{J}(Z_*, K_*)) \leq \sum_{k,l=1}^{K} \mathbb{P}(j \in J^*_{l,k} \cap (\breve{J}_{l,k})^c) \leq$$
$$K_*^2 \exp\left\{ -2 \tilde{C}_0^{\,2} n K_*^{-2}\, \varpi^2(n, K_*) \right\} \leq K_*^2 n^{-1} e^{-t} \leq e^{-t}$$

which completes the proof.

Proof of Lemma 4.2.2. Let us prove the lemma by contradiction. Assume that (4.12) does not holds and

$$\breve{J}_{k,l}(\hat{Z}_K, K) \subset \hat{J}_{k,l}(\hat{Z}_K, K) \tag{4.59}$$

Note that, under the condition (4.59), one has

$$A^{(k,l)}(\hat{Z}_K, \hat{K}) = \Pi_{\breve{J}(k,l)} \left(A^{(k,l)}(\hat{Z}_K, \hat{K}) \right) = \Pi_{\hat{J}(k,l)} \left(A^{(k,l)}(\hat{Z}_K, \hat{K}) \right)$$

so that

$$\|A^{(k,l)}(\hat{Z}_K, \hat{K}) - \Pi_{(1)} \left(\Pi_{\breve{J}(k,l)} \left(A^{(k,l)}(\hat{Z}_K, \hat{K}) \right) \right) \|_F^2 = \|A^{(k,l)}(\hat{Z}_K, \hat{K}) - \Pi_{(1)} \left(\Pi_{\hat{J}(k,l)} \left(A^{(k,l)}(\hat{Z}_K, \hat{K}) \right) \right) \|_F^2$$

Hence (4.7) and (4.59) imply that $\text{Pen}(n, \hat{J}, \hat{K}) \leq \text{Pen}(n, \breve{J}, \hat{K})$. Under assumption (4.11), the latter leads to

$$\mathscr{F}(|\hat{J}_{k,l}|, n_k) + \mathscr{F}(|\hat{J}_{l,k}|, n_l) \leq \mathscr{F}(|\breve{J}_{k,l}|, n_k) + \mathscr{F}(|\breve{J}_{l,k}|, n_l)$$

which contradicts (4.59). In order to complete the proof, apply inequality (4.4).

Proof of Lemma 4.3.1. Note that the difference between separable and non-separable penalty is given by

$$\Delta^{n/s} = \text{Pen}^{(ns)}(n, J, K) - \text{Pen}^{(s)}(n, J, K) = \beta_1 \Delta_1^{n/s} + \beta_2 \Delta_2^{n/s} \tag{4.60}$$

where

$$\Delta_1^{n/s} = |J| \ln \left(\frac{nKe}{|J|} \right) - \sum_{k,l=1}^{K} |J_{k,l}| \ln \left(\frac{n_k e}{|J_{k,l}|} \right), \quad \Delta_2^{n/s} = 2 \ln n - K \sum_{k=1}^{K} \ln n_k.$$

Note that, due to the log-sum inequality (Theorem 17.1.2 of Cover and Thomas (2006)), $\Delta_1^{n/s} \leq 0$ with $\Delta_2^{n/s} = 0$ if and only if $n_k/|J_{k,l}| = nK/|J|$ for every $k, l = 1, \ldots, K$. In the extreme case where the nodes have nonzero connection probabilities only to the nodes in the same class, one has

$|J_{k,l}| = n_k$ for $k = l$ and 0 otherwise, so that $|J| = n$. Then, $\Delta_1^{n/s} = n \ln K$, so that

$$0 \le \Delta_1^{n/s} \le n \ln K. \tag{4.61}$$

Now, consider $\Delta_2^{n/s}$. Note that application of the log-sum inequality (Theorem 17.1.2 of Cover and Thomas (2006)) yields

$$2 \ln n - K^2 \ln(n/K) \le \Delta_2^{n/s} \le 2 \ln n - K \ln(n + 1 - K).$$

It is easy to see that $0 < K^2 \ln n \le n \ln K$ if $n \ge 8$ and $K \le \sqrt{n/ \ln n}$, therefore,

$$2 \ln n - n \ln K \le \Delta_2^{n/s} \le 2 \ln n. \tag{4.62}$$

Combining (4.60)– (4.62), obtain that

$$\beta_2(2 \ln n - n \ln K) \le \Delta^{n/s} \le \beta_1 n \ln K + 2\, \beta_2 \ln n.$$

Hence,

$$\text{Pen}^{(ns)}(n, J, K) \le \text{Pen}^{(s)}(n, J, K) + \beta_1 n \ln K + 2\, \beta_2 \ln n < (2 + \beta_1/\beta_2)\text{Pen}^{(s)}(n, J, K)$$
$$\text{Pen}^{(s)}(n, J, K) \le \text{Pen}^{(ns)}(n, J, K) + \beta_2(2 \ln n - n \ln K) < 2\text{Pen}^{(ns)}(n, J, K),$$

which leads to (4.17).

4.4.4 Supplementary Lemmas

Lemma 4.4.1. *Let A and B be arbitrary matrices in $\mathbb{R}^{m\times n}$ and $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$ be any unit vectors. Let $\tilde{u}, \tilde{v}$ be the singular vectors of matrix A corresponding to its largest singular value. Then,*

$$\langle \Pi_{u,v}(B), A - \Pi_{u,v}(A)\rangle = 0 \quad \text{and} \quad \|A - \Pi_{\tilde{u},\tilde{v}}(A)\| \le \|A - \Pi_{u,v}(A)\|, \tag{4.63}$$

so that, the best rank one approximation of A is given by $\Pi_{(1)}(A) = \Pi_{\tilde{u},\tilde{v}}(A)$. Here, $\Pi_{u,v}(A)$ is defined in (3.5).

Lemma 4.4.2. *Let $A = P + \Xi$. Denote by $(\hat{u}, \hat{v})$ and (u, v) the pairs of singular vectors of matrices $\Pi_J(A)$ and $\Pi_J(P)$, respectively, corresponding to their largest singular values. Then,*

$$\|\Pi_{u,v}(\Pi_J(P)) - P\|_F \le \|\Pi_{\hat{u},\hat{v}}(\Pi_J(P)) - P\|_F \le \|\Pi_{\hat{u},\hat{v}}(\Pi_J(A)) - P\|_F \tag{4.64}$$

where, for any matrix X, $\Pi_{u,v}(X)$ is the projection of X onto the pair of unit vectors (u, v), given in (3.5)*, and $\Pi_J(X)$ is the projection of the matrix X onto the set of all matrices with the rectangular support J.*

Proof. Note that

$$\|\Pi_{\hat{u},\hat{v}}(\Pi_J(A)) - P\|_F^2 = \|\Pi_{\hat{u},\hat{v}}(\Pi_J(P + \Xi)) - P\|_F^2 =$$
$$\|\Pi_{\hat{u},\hat{v}}(\Pi_J(P)) + \Pi_{\hat{u},\hat{v}}(\Pi_J(\Xi)) - P\|_F^2 =$$
$$\|\Pi_{\hat{u},\hat{v}}(\Pi_J(\Xi)) + [\Pi_{\hat{u},\hat{v}}(\Pi_J(P)) - \Pi_J(P)] + [\Pi_J(P) - P]\|_F^2$$

Since matrices $\Pi_{\hat{u},\hat{v}}(\Pi_J(\Xi))$ and $[\Pi_{\hat{u},\hat{v}}(\Pi_J(P)) - \Pi_J(P)]$ are supported on the set of indices J and $\Pi_J(P) - P$ is supported on J^c, the latter matrix is orthogonal to the first two. On the other hand,

$\Pi_{\hat{u},\hat{v}}(\Pi_J(\Xi))$ and $[\Pi_{\hat{u},\hat{v}}(\Pi_J(P)) - \Pi_J(P)] = \Pi^{\perp}_{\hat{u},\hat{v}}(\Pi_J(P))$ are also orthogonal. Therefore,

$$\|\Pi_{\hat{u},\hat{v}}(\Pi_J(A)) - P\|_F^2 = \|\Pi_{\hat{u},\hat{v}}(\Pi_J(\Xi))\|_F^2 + \|\Pi_{\hat{u},\hat{v}}(\Pi_J(P)) - \Pi_J(P)\|_F^2 + \|\Pi_J(P) - P\|_F^2 =$$

$$\|\Pi_{\hat{u},\hat{v}}(\Pi_J(\Xi))\|_F^2 + \|\Pi_{\hat{u},\hat{v}}(\Pi_J(P)) - P\|_F^2 \geq \|\Pi_{\hat{u},\hat{v}}(\Pi_J(P)) - P\|_F^2 \geq \|\Pi_{u,v}(\Pi_J(P)) - P\|_F^2$$

where the last inequality follows from Lemma 4.4.1.

Lemma 4.4.3. *Let elements of matrix $\Xi \in (-1,1)^{n\times n}$ be independent Bernoulli errors. Let matrix Ξ be partitioned into K^2 sub-matrices $\Xi^{(k,l)}$ with supports $J^{(k,l)} = J_{k,l} \times J_{l,k}$, $k,l = 1,\cdots,K$, such that $\Xi^{(k,l)} = (\Xi^{(l,k)})^T$. Then, for any $x > 0$*

$$\mathbb{P}\left\{\sum_{k,l=1}^{K} \left\|\Pi_{J^{(k,l)}}\left(\Xi^{(k,l)}\right)\right\|_{op}^2 \leq C_1|J| + C_2 x\right\} \geq 1 - \exp(-x), \tag{4.65}$$

where C_1 and C_2 are absolute constants independent of n, K and sets $J_{k,l}$, $k,l = 1,\cdots,K$.

Proof. Denote $|J_{k,l}| = n_{k,l}$, $k,l = 1,\cdots,K$, and observe that matrices $\Xi^{(k,l)}$ are effectively of the size $n_{k,l} \times n_{l,k}$. Consider $K(K+1)/2$-dimensional vectors ξ and μ with elements $\xi_{k,l} = \|\Pi_{J^{(k,l)}}\left(\Xi^{(k,l)}\right)\|_{op}$ and $\mu_{k,l} = \mathbb{E}\|\Pi_{J^{(k,l)}}\left(\Xi^{(k,l)}\right)\|_{op}$, $1 \leq k \leq l \leq K$, and let $\eta = \xi - \mu$. Then,

$$\Delta = \sum_{k,l=1}^{K} \left\|\Pi_{J^{(k,l)}}\left(\Xi^{(k,l)}\right)\right\|_{op}^2 \leq \|\xi\|^2 \leq 2\|\eta\|^2 + 2\|\mu\|^2 \tag{4.66}$$

Hence, we need to construct the upper bounds for $\|\eta\|^2$ and $\|\mu\|^2$.

We start with constructing upper bounds for $\|\mu\|^2$. Let $\Xi_{i,j}^{(k,l)}$ be elements of the $(n_{k,l} \times n_{l,k})$-dimensional matrix $\Pi_{J^{(k,l)}}\left(\Xi^{(k,l)}\right)$. Then, $\mathbb{E}(\Xi_{i,j}^{(k,l)}) = 0$ and, by Hoeffding's inequality, $\mathbb{E}\left\{\exp(\lambda \Xi_{i,j}^{(k,l)})\right\} \leq \exp\left(\lambda^2/8\right)$. Taking into account that Bernoulli errors are bounded by one in absolute value and applying Corollary 3.3 of Bandeira and van Handel (2016) with $m = n_{k,l}$, $n = n_{l,k}$, $\sigma_* = 1$, $\sigma_1 = \sqrt{n_{l,k}}$ and $\sigma_2 = \sqrt{n_{k,l}}$, obtain

$$\mu_{k,l} \le C_0 \left(\sqrt{n_{k,l}} + \sqrt{n_{l,k}} + \sqrt{\ln(n_{k,l} \wedge n_{l,k})} \right)$$

where C_0 is an absolute constant independent of $n_{k,l}$ and $n_{l,k}$. Therefore,

$$\|\mu\|^2 \le 3C_0^2 \sum_{k,l=1}^{K} (n_{k,l} + n_{l,k} + \ln(n_{k,l} \wedge n_{l,k})) \le 6C_0^2 |J| + 3C_0^2 \sum_{k,l=1}^{K} \ln(n_{k,l}). \tag{4.67}$$

Next, we show that, for any fixed partition, $\eta_{k,l} = \xi_{k,l} - \mu_{k,l}$ are independent sub-gaussian random variables when $1 \le k \le l \le K$. Independence follows from the conditions of Lemma 4.4.3. To prove the sub-gaussian property, use Talagrand's concentration inequality (Theorem 6.10 of Boucheron et al. (2013)): if $\Xi_1, \Xi_2, \Xi_3, \cdots, \Xi_n$ are independent random variables taking values in the interval $[0, 1]$ and $f : [0, 1]^n \to R$ is a separately convex function such that $|f(x) - f(y)| \le \|x - y\|$ for all $x, y \in [0, 1]^n$, then, for $Z = f(\Xi_1, \Xi_2, \Xi_3, \cdots, \Xi_n)$ and any $t > 0$, one has $\mathbb{P}(Z > \mathbb{E}Z + t) \le \exp(-t^2/2)$. Apply this theorem to vectors $\zeta_{k,l} = \text{vec}(\Pi_{J(k,l)} \left(\Xi^{(k,l)} \right)) \in [0, 1]^{n_{k,l} \times n_{l,k}}$ and $f(\Pi_{J(k,l)} \left(\Xi^{(k,l)} \right)) = f(\zeta_{k,l}) = \left\| \Pi_{J(k,l)} \left(\Xi^{(k,l)} \right) \right\|_{op}$. Note that, for any two matrices Ξ and $\tilde{\Xi}$ of the same size, one has $\|\Xi - \tilde{\Xi}\|_{op}^2 \le \|\Xi - \tilde{\Xi}\|_F^2 = \|\text{vec}(\Xi) - \text{vec}(\tilde{\Xi})\|^2$. Then, applying Talagrand's inequality with $Z = \|\Pi_{J(k,l)} \left(\Xi^{(k,l)} \right)\|_{op}$ and $Z = -\|\Pi_{J(k,l)} \left(\Xi^{(k,l)} \right)\|_{op}$, obtain

$$\mathbb{P} \left(\left| \|\Pi_{J(k,l)} \left(\Xi^{(k,l)} \right)\|_{op} - \mathbb{E}\|\Pi_{J(k,l)} \left(\Xi^{(k,l)} \right)\|_{op} \right| > t \right) \le 2\exp(-t^2/2).$$

Now, use the Lemma 5.5 of Vershynin (2012) which states that the latter implies that, for any $t > 0$ and some absolute constant $C_4 > 0$,

$$\mathbb{E}\left[\exp(t\eta_{k,l}) \right] = \mathbb{E}\left[\exp(t(\xi_{k,l} - \mu_{k,l})) \right] \le \exp(C_4 t^2/2). \tag{4.68}$$

Hence, $\eta_{k,l}$ are independent sub-gaussian random variables when $1 \leq k \leq l \leq K$.

In order to obtain an upper bound for $\|\eta\|^2$, use Theorem 2.1 of Hsu et al. (2012). Applying this theorem with $A = I_{K(K+1)/2}$, $\mu = 0$ and $\sigma^2 = C_4$ to a sub-vector $\tilde{\eta}$ of η which contains components $\eta_{k,l}$ with $1 \leq k \leq l \leq K$, obtain

$$\mathbb{P}\left\{\|\tilde{\eta}\|^2 \geq C_4 \left(K(K+1)/2 + \sqrt{2\,K(K+1)\,x} + 2x\right)\right\} \leq \exp(-x).$$

Since $\|\eta\|^2 \leq 2\|\tilde{\eta}\|^2$, derive

$$\mathbb{P}\left\{\|\eta\|^2 \geq 2C_4K(K+1) + 6C_4x\right\} \leq \exp\left(-x\right) \tag{4.69}$$

Combination of formulas (4.66) and (4.69) yield

$$\mathbb{P}\left\{\|\xi\|^2 \leq 2\|\mu\|^2 + 4C_4K(K+1) + 12C_4x\right\} \geq 1 - \exp\left(-x\right)$$

Plugging in $\|\mu\|^2$ from (4.67) into the last inequality, derive for any $x > 0$ that

$$\mathbb{P}\left\{\|\xi\|^2 \leq 12C_0^2|J| + 6C_0^2 \sum_{k,l=1}^{K} \ln(n_{k,l}) + 4C_4K(K+1) + 12C_4x\right\} \geq 1 - \exp\left(-x\right). \tag{4.70}$$

Since $K(K+1) \leq 2K^2$ and

$$6C_0^2 \sum_{k,l=1}^{K} \ln(n_{k,l}) + 8C_4K^2 \leq \max(6C_0^2, 8C_4) \sum_{k,l=1}^{K} \ln(n_{k,l}e) \leq \max(6C_0^2, 8C_4)|J|,$$

inequality (4.65) holds with $C_1 = 12C_0^2 + \max(6C_0^2, 8C_4)$ and $C_2 = 12C_4$.

Lemma 4.4.4. *For any $t > 0$,*

$$\mathbb{P}\left\{\sum_{k,l=1}^{\hat{K}}\left\|\Pi_{\hat{J}(k,l)}\left(\Xi^{(k,l)}(\hat{Z},\hat{K})\right)\right\|_{op}^{2} - F_1(n,\hat{J},\hat{K}) \le C_2 t\right\} \ge 1-\exp(-t), \tag{4.71}$$

with $F_1(n,J,K) = F_1^{(ns)}(n,J,K)$ given by (4.26)*, or $F_1(n,J,K) = F_1^{(s)}(n,J,K)$ given by* (4.27).

Proof. Note that $|J_{k,l}| \le |J_{k,l}|\ln(nKe/|J_{k,l}|)$, $|J| \le |J|\ln(nKe/|J|)$, and also that $|J| = \sum_{k,l=1}^{K}|J_{k,l}|$. First, let us prove the statement for $F_1(n,J,K) = F_1^{(ns)}(n,J,K)$. For this purpose, set $x = t + 3\ln n + n\ln K + |J|\ln(nKe/|J|)$ in Lemma 4.4.3 and apply the union bound over $K \in [1,n]$, $Z \in \mathcal{M}_{n,K}$ and $J \subseteq \{1,\ldots,nK\}$. Obtain

$$\begin{aligned}
&\mathbb{P}\left\{\sum_{k,l=1}^{\hat{K}}\left\|\Pi_{\hat{J}(k,l)}\left(\Xi^{(k,l)}(\hat{Z},\hat{K})\right)\right\|_{op}^{2} - F_1^{(ns)}(n,\hat{J},\hat{K}) - C_2 t \ge 0\right\}\\
&\le \sum_{K=1}^{n}\sum_{Z\in\mathcal{M}_{n,K}}\sum_{j=1}^{nK}\sum_{|J|=j}\mathbb{P}\left\{\sum_{k,l=1}^{K}\|\Pi_{J(k,l)}\left(\Xi^{(k,l)}(Z,K)\right)\|_{op}^{2} - F_1^{(ns)}(n,J,K) \ge C_2 t\right\}\\
&\le \sum_{K=1}^{n}\sum_{Z\in\mathcal{M}_{n,K}}\sum_{j=1}^{nK}\sum_{|J|=j}\exp(-t-3\ln n - n\ln K - j\ln(nKe/j))\\
&\le \sum_{K=1}^{n}\sum_{j=1}^{nK}K^n\binom{nK}{j}\exp(-t-3\ln n - n\ln K - j\ln(nKe/j)) \le \exp(-t).
\end{aligned}$$

In order to prove the statement for $F_1(n,J,K) = F_1^{(s)}(n,J,K)$, choose

$$x = t + \ln n + n\ln K + \sum_{k,l=1}^{K}\left[\ln(n_k) + |J_{k,l}|\ln(n_k\, e/|J_{k,l}|)\right]$$

in Lemma 4.4.3 and again apply the union bound over $Z \in \mathcal{M}_{n,K}$, $K \in [1,n]$ and $|J_{kl}| \in$

$\{1,\ldots,n_k\}$, $k,l=1,\ldots,K$. Obtain

$$\begin{aligned}
&\mathbb{P}\left\{\sum_{k,l=1}^{\hat{K}}\left\|\Pi_{\hat{J}(k,l)}\left(\Xi^{(k,l)}(\hat{Z},\hat{K})\right)\right\|_{op}^{2}-F_1^{(s)}(n,\hat{J},\hat{K})-C_2 t\geq 0\right\}\\
&\leq \sum_{K=1}^{n}\sum_{Z\in\mathcal{M}_{n,K}}\prod_{k,l=1}^{K}\sum_{j_{k,l}=1}^{n_k}\sum_{|J_{k,l}|=j_{k,l}}\mathbb{P}\left\{\sum_{k,l=1}^{K}\|\Pi_{J(k,l)}\left(\Xi^{(k,l)}(Z,K)\right)\|_{op}^{2}-F_1^{(s)}(n,J,K)\geq C_2 t\right\}\\
&\leq \sum_{K=1}^{n}K^n\prod_{k,l=1}^{K}\sum_{j_{k,l}=1}^{n_k}\binom{n_k}{j_{k,l}}\exp\left(-t-\ln n-n\ln K-\sum_{k,l=1}^{K}\left[\ln(n_k)+j_{k,l}\ln(n_k\, e/j_{k,l})\right]\right)\\
&\leq \exp(-t),
\end{aligned}$$

which completes the proof.

CHAPTER 5: DISCUSSION AND FUTURE WORK

In this dissertation, we studied the statistical network models with community structure. We reviewed the SBM, DCBM and PABM model. We carry out the in-depth study of the PABM model. Since the SBM and the DCBM are the special cases of the PABM model, the PABM model is more general and flexible model comparison to the existing block models. In the rearranged probability matrix $P(Z, K)$ of the PABM model given by (3.3), we observe that the probability matrix consists of K^2 arbitrary rank one blocks. This demonstrates that rank$(P(Z, K)) =$ rank(P) can take any value between K and K^2. In comparison, all other block models restrict the rank of P to be exactly K. This is true not only for the SBM and the DCBM discussed above but also for their generalizations such as the Mixed Membership models (MMM) (see, e.g., Airoldi et al. (2008) and Cheng et al. (2017)) and the Degree Corrected Mixed Membership (DCMM) (see, e.g., Jin et al. (2017)). While the MMM and the DCMM allows more diverse structures of rank K matrices (those matrices have to be just a product of two rank K matrices with nonnegative components while the PABM require to be a combination of K^2 rank one matrices), meaningful fits of the MMM and DCMM rely on a variety of conditions (one needs to have pure nodes in the network and some identifiability conditions need to be satisfied). In addition, while the MMM and DCMM are extremely useful for analysis of social and society-related networks such as publications networks, they may not be appropriate in some other applications where each node can belong to one and only one class. The butterfly similarity network studied in our paper provides an example of such application. However, while the PABM model is extremely valuable, the statistical inference in Sengupta and Chen (2018) has been incomplete. In particular, the authors considered only the case of a small finite number of communities K; they provided only asymptotic consistency results as $n \to \infty$ without any error bounds when n is finite; their NP-hard clustering procedure was tailored to the case of a small K. In addition, the relaxation of this NP-hard procedure seems to be operational only in the case of $K = 2$ since all simulations and real data examples in Sengupta and Chen (2018) only tackled the case of $K = 2$.

We addressed some of those deficiencies and advance the theory of the PABM. Specifically, the main achievement of our work lies in the fact that, unlike Sengupta and Chen (2018) who worked in terms of maximizing the Poisson likelihood and likelihood modularity, we recognize that the probability matrix of the PABM is formed by a unique collection of rank one matrices. This observation leads to a variety of breakthroughs. In particular, we are able to carry out estimation and clustering for the PABM, without imposing any identifiability conditions, similarly to SBM and unlike the DCBM and mixed membership models. Our understanding of the probability matrix structure leads to the Frobenius norm minimization as the basis of optimization procedure and to estimation of probability matrices by rank one approximations of the community matrices. The latter allows us to deal with the situation when the number of communities is uncertain and is possibly growing with n. Moreover, we are able to derive non-asymptotic upper bounds for the estimation error even in the case when the number of communities is unknown. In addition, we use the accuracy of approximation of the adjacency matrix for various number of communities, to identify the number of communities in the network. Furthermore, we note that, under simple conditions, the columns of the probability matrix that correspond to any of the communities lie in a K-dimensional subspace which is different from subspaces corresponding to all other communities. The latter conclusion results in the introduction of the Sparse Subspace Clustering (SSC) approach for partitioning the network into communities.

The real life networks are usually sparse in a sense that a large number of nodes have small degrees. One of the advantages of the PABM is that it allows flexible modeling of sparsity. Traditionally, in most statistical models, sparsity of a vector means that a large proportion of its components is equal to zero. One of the shortcomings of both the SBM and the DCBM is that they do not allow to impose the condition that some of the connection probabilities are equal to zero. Indeed, for the SBM, it is not realistic to assume that all nodes in a pair of communities have no connections. Neither can one set any of the node-specific weight to zero, since this will force the respective node to be totally disconnected from the network. For this reason, unlike in other numerous statistical settings, sparsity in block models is defined as a low maximum probability of connections between the nodes: : $\max_{i,j} P_{i,j} \leq \rho(n)$ where $\rho(n)$ is small when n is large (see, e.g., Klopp, Tsybakov and Verzelen (2017), Lei and Rinaldo (2015)). In order take advantage of this definition of sparsity, even in the simplest model, the SBM, one needs to carry out the estimation under the restriction that all entries of the matrix $\hat{P}$ are bounded above by $\rho(n)$ (see Klopp, Tsybakov and Verzelen (2017)). In addition, this definition prevents nodes from having high degrees. On the contrary, the PABM setting allows some connection probabilities to be zero while keeping average connection probabilities between classes above certain level and the network connected. Indeed, in the context of PABM, setting $\Lambda_i^{(k,l)} = 0$ simply means that that node i in class k is not active ("popular") in class l. This, nevertheless, does not prevent this node from having high probability of connection with nodes in another class. Similarly, to other sparse statistical settings, allowing structural sparsity (i.e., setting connection probabilities to zero rather than to a very small positive number) not only leads to better understanding of network topology but leads to more precise estimation of the probability matrix P_*.

We carry out the in depth investigation of the sparse PABM model. Before starting the estimation procedure, we introduced the sparsity sets J_* consisting the set of all non zero indices of the true popularity matrix Λ_*. We imposed the penalty on the sparsity set since the estimator of the sparsity set J_* by the support set of the adjacency matrix is not always be best estimator. As we shown in example, sometimes the zero value in the probability matrix is estimated by the non zero values. We introduced the penalty that will lead to trimming the support of the estimated probability matrix. We considered the separable and non separable penalty. The separable penalty has a property that the support of the adjacency matrix is contained in the support of the probability matrix. We further showed that for a network with balanced community if the nonzero entries in the true probability matrix is above some threshold, then the support of the probability matrix and adjacency matrix coincide with high probability. The non separable penalty in the other hand is easy to interpret since it is expressed in terms of the support sets of probability matrix rather than the supports of the individual blocks in the separable case. We showed that both penalties are within constant factor of each other. We estimated the probability matrix and it's true support using Frobenius norm minimization procedure. This estimator is shown to be better than the one in regular PABM in the sparse network which is demonstrated by simulation and real data examples in our paper Noroozi et al. (2019*b*).

In this dissertation, we worked on the network modeling of static network. There are many networks that change in time. For instance, one can consider social networks, such as Facebook or Twitter, in which the group dynamics changes with time, as users frequently join and leave the groups. For this reason, we consider the extension of our research to dynamic setting as a matter of future work.

In addition, in the context of the modern era of big data, there is a need for the analysis of collections of network data objects. For example: in brain connectomics studies, a sample of networks is available form multiple populations of interests such as ill patients and healthy controls. These types of data are growing in the fields of systems biology, neuroscience, etc. This motivates studying of the new class of models consisting of observations of multiple networks. This is another direction of our future research.

LIST OF REFERENCES

Airoldi, E. M., Blei, D. M., Fienberg, S. E. and Xing, E. P. (2008), 'Mixed membership stochastic blockmodels', *J. Mach. Learn. Res.* **9**, 1981–2014.

Amini, A. A., Chen, A., Bickel, P. J. and Levina, E. (2013), 'Pseudo-likelihood methods for community detection in large sparse networks', *Ann. Statist.* **41**(4), 2097–2122.
URL: *https://doi.org/10.1214/13-AOS1138*

Bandeira, A. S. and van Handel, R. (2016), 'Sharp nonasymptotic bounds on the norm of random matrices with independent entries', *The Annals of Probability* **44**(4), 2479–2506.

Barabási, A.-L. et al. (2016), *Network science*, Cambridge university press.

Berkson, J. (1950), 'Are there two regressions?', *Journal of the american statistical association* **45**(250), 164–180.

Boucheron, S., Lugosi, G. and Massart, P. (2013), 'Concentration inequalities: A nonasymptotic theory of independence'.

Bovy, J., Hennawi, J. F., Hogg, D. W., Myers, A. D., Kirkpatrick, J. A., Schlegel, D. J., Ross, N. P., Sheldon, E. S., McGreer, I. D., Schneider, D. P. and Weaver, B. A. (2011), 'Think outside the color box: Probabilistic target selection and the sdss-xdqso quasar targeting catalog', *Astrophysical Journal* **729**(2).

Carroll, R. J., Delaigle, A. and Hall, P. (2009), 'Nonparametric prediction in measurement error models', *Journal of the American Statistical Association* **104**(487), 993–1003.

Carroll, R. J., Ruppert, D., Crainiceanu, C. M. and Stefanski, L. A. (2006), *Measurement error in nonlinear models: a modern perspective*, Chapman and Hall/CRC.

Celisse, A., Daudin, J.-J., Pierre, L. et al. (2012), 'Consistency of maximum-likelihood and variational estimators in the stochastic block model', *Electronic Journal of Statistics* **6**, 1847–1899.

Cheng, J., Li, T., Levina, E. and Zhu, J. (2017), 'High-dimensional mixed graphical models', *Journal of Computational and Graphical Statistics* **26**(2), 367–378.

Comte, F. and Kappus, J. (2015), 'Density deconvolution from repeated measurements without symmetry assumption on the errors', *Journal of Multivariate Analysis* **140**, 31–46.

Comte, F. and Lacour, C. (2011), 'Data-driven density estimation in the presence of additive noise with unknown distribution', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **73**(4), 601–627.

Cover, T. M. and Thomas, J. A. (2006), *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*, Wiley-Interscience, New York, NY, USA.

Davidson, K. R. and Szarek, S. J. (2001), Chapter 8 - local operator theory, random matrices and banach spaces, *in* W. Johnson and J. Lindenstrauss, eds, 'Handbook of the Geometry of Banach Spaces', Vol. 1 of *Handbook of the Geometry of Banach Spaces*, Elsevier Science B.V., pp. 317 – 366.

URL: *http://www.sciencedirect.com/science/article/pii/S1874584901800103*

Delaigle, A. (2007), 'Nonparametric density estimation from data with a mixture of berkson and classical errors', *Canadian Journal of Statistics* **35**(1), 89–104.

Delaigle, A. (2008), 'An alternative view of the deconvolution problem', *Statistica Sinica* **18**(3), 1025–1045.

Du, L., Zou, C. and Wang, Z. (2011), 'Nonparametric regression function estimation for errors-in-variables models with validation data', *Statistica Sinica* pp. 1093–1113.

E. Fienberg, S., M. Meyer, M. and Wasserman, S. (1985), ‘Statistical analysis of multiple sociometric relations’, *Journal of The American Statistical Association - J AMER STATIST ASSN* **80**, 51–67.

Geng, P. and Koul, H. L. (2018), ‘Minimum distance model checking in berkson measurement error models with validation data’, *TEST* pp. 1–21.

Giraud, C. (2015), *Introduction to high-dimensional statistics*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability, CRC Press, Hoboken, NJ.

Goldenberg, A., Zheng, A. X., Fienberg, S. E. and Airoldi, E. M. (2010), ‘A survey of statistical network models’, *Foundations and Trends® in Machine Learning* **2**(2), 129–233.

Goldenshluger, A. (1999), ‘On pointwise adaptive nonparametric deconvolution’, *Bernoulli* **5**(5), 907–925.
URL: *https://projecteuclid.org:443/euclid.bj/1171290404*

Holland, P. W., Laskey, K. B. and Leinhardt, S. (1983), ‘Stochastic blockmodels: First steps’, *Social Networks* **5**(2), 109 – 137.
URL: *http://www.sciencedirect.com/science/article/pii/0378873383900217*

Hsu, D., Kakade, S. and Zhang, T. (2012), ‘A tail inequality for quadratic forms of subgaussian random vectors’, *Electron. Commun. Probab.* **17**, 6 pp.
URL: *https://doi.org/10.1214/ECP.v17-2079*

Jin, J., Ke, Z. T. and Luo, S. (2017), ‘Estimating network memberships by simplex vertex hunting’, *arXiv e-prints* p. arXiv:1708.07852.

Karrer, B. and Newman, M. (2011), ‘Stochastic blockmodels and community structure in networks’, *Physical review. E, Statistical, nonlinear, and soft matter physics* **83**, 016107.

Kim, K. H., Chao, S.-K. and Härdle, W. K. (2016), 'Simultaneous inference for the partially linear model with a multivariate unknown function when the covariates are measured with errors'.

Klopp, O., Lounici, K. and Tsybakov, A. B. (2017), 'Robust matrix completion', *Probability Theory and Related Fields* **169**(1), 523–564.

Klopp, O., Tsybakov, A. B. and Verzelen, N. (2017), 'Oracle inequalities for network models and sparse graphon estimation', *Ann. Statist.* **45**(1), 316–354.

Lei, J. and Rinaldo, A. (2015), 'Consistency of spectral clustering in stochastic block models', *Ann. Statist.* **43**(1), 215–237.

Long, J. P., El Karoui, N. and Rice, J. A. (2016), 'Kernel density estimation with berkson error', *Canadian Journal of Statistics* **44**(2), 142–160.

Lorrain, F. and White, H. C. (1971), 'Structural equivalence of individuals in social networks', *The Journal of Mathematical Sociology* **1**(1), 49–80.
URL: *https://doi.org/10.1080/0022250X.1971.9989788*

Lust-Piquard, F. and Pisier, G. (1991), 'Non commutative khintchine and paley inequalities', *Ark. Mat.* **29**(1-2), 241–260.
URL: *https://doi.org/10.1007/BF02384340*

Meister, A. (2009), *Deconvolution Problems in Nonparametric Statistics*, Vol. 193, Springer Publishing Company, Incorporated.

Noroozi, M., Rimal, R. and Pensky, M. (2019*a*), 'Estimation and Clustering in Popularity Adjusted Stochastic Block Model', *arXiv e-prints* p. arXiv:1902.00431.

Noroozi, M., Rimal, R. and Pensky, M. (2019*b*), 'Sparse popularity adjusted stochastic block model'.

Rimal, R. and Pensky, M. (2019), ‘Density deconvolution with small berkson errors’, *Mathematical Methods of Statistics* **28**, 208 – 227.
URL: *https://doi.org/10.3103/S1066530719030025*

Robinson, E. (1999), ‘Seismic inversion and deconvolution’, *Elsevier, Oxford* .

Rohe, K., Chatterjee, S. and Yu, B. (2011), ‘Spectral clustering and the high-dimensional stochastic blockmodel’, *Ann. Statist.* **39**(4), 1878–1915.
URL: *https://doi.org/10.1214/11-AOS887*

Rudelson, M. and Vershynin, R. (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’, World Scientific, pp. 1576–1602.

Sengupta, S. and Chen, Y. (2018), ‘A block model for node popularity in networks with community structure’, *Journal of the Royal Statistical Society Series B* **80**(2), 365–386.

Silverman, B. W. (1986), *Density estimation for statistics and data analysis / B.W. Silverman*, Chapman and Hall London ; New York.

Soltanolkotabi, M. and Candes, E. J. (2012), ‘A geometric analysis of subspace clustering with outliers’, *Ann. Statist.* **40**(4), 2195–2238.

Talagrand, M. (2014), *Upper and lower bounds for stochastic processes: modern methods and classical problems*, Vol. 60, Springer Science & Business Media.

Vershynin, R. (2012), *Introduction to the non-asymptotic analysis of random matrices*, Cambridge University Press, p. 210–268.

Vershynin, R. (2018), *High-Dimensional Probability: An Introduction with Applications in Data Science*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.

Wang, L. (2003), 'Estimation of nonlinear berkson-type measurement error models', *Statistica Sinica* pp. 1201–1210.

Wang, L. et al. (2004), 'Estimation of nonlinear models with berkson measurement errors', *The Annals of Statistics* **32**(6), 2559–2579.

Wason, C. B., Black, J. and King, G. (1984), 'Seismic modeling and inversion', *Proceedings of the IEEE* **72**(10), 1385–1393.