

Fundamental Limits of Spectral Clustering in Stochastic Block Models

Anderson Ye Zhang^{ID}

Abstract—Spectral clustering has been widely used for community detection in network sciences. While its empirical successes are well-documented, a clear theoretical understanding, particularly for sparse networks where degrees are much smaller than $\log n$, remains unclear. In this paper, we address this significant gap by demonstrating that spectral clustering offers exponentially small error rates when applied to sparse networks under Stochastic Block Models. Our analysis provides sharp characterizations of its performance, backed by matching upper and lower bounds possessing an identical exponent with the same leading constant. The key to our results is a novel truncated ℓ_2 perturbation analysis for eigenvectors, coupled with a new analysis idea of eigenvectors truncation.

Index Terms—Spectral clustering, community detection, stochastic block model, spectral perturbation, network analysis.

I. INTRODUCTION

COMMUNITY detection [1], [2], [3] is a central problem in network science. The goal is to recover hidden community structures from network data and has broad applications in social science, neuroscience, computer science, and physics. Among various approaches for community detection, spectral clustering [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15] is a particularly popular one and has achieved tremendous success. It first reduces the dimensionality of data by a spectral decomposition and performs clustering in a reduced-dimension space. It is computationally appealing, easy to implement, and has surprisingly good performance.

Driven by its popularity and success, there has been growing interest in theoretical and statistical analysis for the performance of spectral clustering. In most literature [4], [16], [17], [18], [19], spectral clustering is shown to attain polynomially small errors. As a result, it is often used as a warm start to initialize [19], [20], [21], [22], [23] delicate procedures in order to provably achieve exponentially small errors and even optimal statistical accuracy. However, spectral clustering performs exceptionally well numerically, indicating a gap between theory and practice. This raises important questions:

Does spectral clustering achieve exponentially small errors? How small can it be?

These questions are answered in a seminal work [24] that considers a very special setting where the network has two equal-sized communities and $p = a(\log n)/n, q = b(\log n)/n$ with fixed constants a, b . The spectral clustering procedure studied in [24] has a simple form: it only utilizes the second leading eigenvector of the adjacency matrix and partitions the network using signs of coordinates. Under this setting, [24] proves the spectral clustering achieves an optimal exponential rate. However, the procedure, result, and proof technique of [24] are limited to this special setting and cannot be extended to even slightly more general cases especially when p, q are of a smaller order of $(\log n)/n$. This leads to the following open problem: *Can we establish sharp exponential error bound for spectral clustering on sparse networks when the degrees are far smaller than $\log n$?*

In this paper, we address this open problem by demonstrating that spectral clustering offers exponentially small error rates when applied to sparse networks under the Stochastic Block Model (SBM) [25] which is the most studied model for community detection. In addition, we provide a matching lower bound that has the same exponent, including the leading constant, as the upper bound. The matching upper and lower bounds together give a sharp characterization of the performance of spectral clustering, demonstrating its ability and also limit. Hence, we refer to our results as the fundamental limits of spectral clustering in SBMs. We emphasize that this is different from information-theoretical analysis for SBMs, referred to as fundamental limits of SBMs in [1] and minimax rates of SBMs in [26]. See Section III-C for further elaboration.

Consider an n -node network with k communities. For every two nodes, we observe an edge with probability p if they belong to the same community and q otherwise. The goal is to recover the hidden community structure z^* given the network. We study a popular spectral clustering procedure [16], [18], [19], [27] \hat{z} that is based on the eigendecomposition of the adjacency matrix. It first regularizes [28] the network by removing high-degree nodes, a step that is necessary for the concentration of sparse networks [29], [30]. It then weights leading eigenvectors of the regularized adjacency matrix by corresponding eigenvalues, followed by the k -means clustering. Its clustering error can be measured by a loss $\ell(\hat{z}, z^*)$. See Sections II-A-II-B for more details about the model and

Manuscript received 25 October 2023; revised 9 May 2024; accepted 24 June 2024. Date of publication 9 July 2024; date of current version 17 September 2024. This work was supported in part by NSF under Grant DMS-2112988 and in part by the Alfred H. Williams Faculty Scholar Award.

The author is with the Department of Statistics and Data Science, University of Pennsylvania, Philadelphia, PA 19104 USA (e-mail: ayz@wharton.upenn.edu).

Communicated by S. Jalali, Associate Editor for Machine Learning and Statistics.

Digital Object Identifier 10.1109/TIT.2024.3425581

0018-9448 © 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See <https://www.ieee.org/publications/rights/index.html> for more information.

the algorithm. The main result of this paper is summarized below in Theorem 1.

Theorem 1: Assume k is a constant and all community sizes n_1, n_2, \dots, n_k are of the same order. Assume p, q satisfy $0 < q < p \leq 1/10$ and are of the same order. We further assume $\frac{n(p-q)^2}{p} \rightarrow \infty$. Then

$$\begin{aligned} \mathbb{E}\ell(\hat{z}, z^*) &\leq \exp(-(1 - o(1))J_{\min}) + 2n^{-3}, \\ \text{and } \mathbb{E}\ell(\hat{z}, z^*) &\geq \exp(-(1 + o(1))J_{\min}) - 2n^{-3}, \end{aligned}$$

where

$$J_{\min} := \min_{1 \leq a \neq b \leq k} \max_{t > 0} \left((n_a - n_b)t \frac{p+q}{2} - n_a \log(qe^t + 1 - q) - n_b \log(pe^{-t} + 1 - p) \right). \quad (1)$$

Theorem 1 gives both upper and lower bounds for the partial recovery of spectral clustering \hat{z} . We note that the additive term $2n^{-3}$ in the error bounds can be replaced by n^{-C} for an arbitrarily large constant $C > 0$, and in general should be ignored. Then the upper and lower bounds are both in an exponential form with a matching asymptotic exponent J_{\min} . In this way, Theorem 1 gives an exact characterization of the error exponent of the spectral clustering, even including the sharp leading constant. Theorem 1 holds under mild conditions, allowing networks to have multiple imbalanced communities and to be sparse with $p, q \ll (\log n)/n$. The assumption $n(p-q)^2/p \rightarrow \infty$ is known to be the necessary and sufficient condition to have consistent community detection [26]. It is also worth mentioning that Theorem 1 characterizes precisely the performance of the spectral clustering for each instance of z^* , which is beyond the minimax framework that only focuses on the worst case of a large parameter space.

The asymptotic exponent J_{\min} in Theorem 1 has a complicated dependence on p, q and the community sizes. Despite no explicit expression, the quantity J_{\min} is closely related to tail probabilities of Bernoulli random variables. Let $\{X_i\}$ and $\{Y_j\}$ be independent Bernoulli random variables with probabilities q, p respectively. By Chernoff bound (see Lemma 15),

$$\begin{aligned} \min_{1 \leq a \neq b \leq k} -\log \mathbb{P} \left(\sum_{i \in [n_a]} X_i - \sum_{j \in [n_b]} Y_j \geq (n_a - n_b) \frac{p+q}{2} \right) \\ = (1 + o(1))J_{\min}. \end{aligned}$$

To further explain why J_{\min} appears in Theorem 1 and is fundamental for the performance of the spectral clustering, we study an oracle estimator that is inspired by \hat{z} but utilizes the unobserved population eigenstructure instead of the sample one. Its corresponding statistical accuracy turns out to be determined by the aforementioned tail probabilities (see Section II-D).

Under the setting of Theorem 1, J_{\min} and $n(p-q)^2/p$ can be shown to be of the same order (see Lemma 13). Hence, the assumption $n(p-q)^2/p \rightarrow \infty$ in Theorem 1 can be replaced by $J_{\min} \rightarrow \infty$ and is the sufficient and necessary condition for \hat{z} to have a vanishing error. Theorem 1 also immediately indicates a sharp threshold of the exact recovery. When $J_{\min} \geq (1 + \epsilon) \log n$ for any constant $\epsilon > 0$, \hat{z} achieves

the exact recovery (i.e., $\ell(\hat{z}, z^*) = 0$) with high probability. When $J_{\min} \leq (1 - \epsilon) \log n$ for any constant $\epsilon > 0$, \hat{z} fails to achieve the exact recovery with constant probability, i.e., $\mathbb{P}(\ell(\hat{z}, z^*) \neq 0) \geq c$ for some constant $c > 0$.

The most related work in the literature is [24]. As mentioned earlier, [24] considers a very special setting where the network has two equal-sized communities and $p = a(\log n)/n, q = b(\log n)/n$ with fixed constants a, b and its results cannot be extended to even slightly more general cases. On the contrary, our results hold for general SBMs and cover the setting of [24] as a special case. Other closely related papers including [31], [32], [33] obtain exponential error bounds for spectral clustering under Gaussian and sub-Gaussian mixture models. However, they rely heavily on Gaussianity and sub-Gaussianity of data. Direct application of their results to networks or to other binary data can only lead to trivial upper bounds.

The key to Theorem 1 is a novel truncated ℓ_2 perturbation analysis for eigenvectors and a proof idea of eigenvectors truncation. Let u be one of the leading k eigenvectors of the regularized adjacency matrix. Analysis of the spectral clustering shows that tail probabilities of quantities in the form of $\sum_{i \in [n]} u_i(X_i - \mathbb{E}X_i)$ play a crucial role, where $\{X_i\}_{i \in [n]}$ are some Bernoulli random variables such as edges of a node. Direction applications of classical concentration inequalities for Bernoulli random variables often involve $\|u\|_\infty$ that has some inevitable $\log n$ factor. To deal with this $\log n$ factor and to derive meaningful tail probabilities, p has to be at least of an order $(\log n)/n$ as seen in literature [24], [34], [35]. However, in this paper, we consider sparse networks with the connectivity probability allowed to be far smaller than $(\log n)/n$ and the aforementioned analysis breaks down. Instead, we truncate coordinates of u by some carefully selected threshold t_0 . The truncated eigenvector has an ℓ_∞ norm bounded by t_0 and its inner product with $\{X_i - \mathbb{E}X_i\}$ is well-controlled with desired tail probabilities. The approximation error of replacing u by its truncated counterpart turns out to be related to $\sum_{i \in [n]} \|u_i\|^2 \mathbb{I}\{|u_i| \geq t_0\}$, which we refer to as a truncated ℓ_2 norm of u . We establish an upper bound for the truncated ℓ_2 norm in Theorem 3 and further show such approximation error is negligible. The idea of eigenvector truncation and Theorem 3 are critical to establishing sharp upper and lower bounds in Theorem 1, and might be useful for fine-grained spectral perturbation analysis of other binary random matrices. To establish Theorem 1, we also use a leave-one-out technique [24], [36] to decouple dependence between the eigenvectors and the regularized adjacency matrix.

We conclude this section by summarizing contributions of this paper:

- 1) Our result is the first in the literature to show spectral clustering has exponentially small error rate for sparse networks.
- 2) Our characterization of the performance of spectral clustering is sharp and precise. In addition to the upper bound, we provide a matching lower bound that has the same exponent, including the leading constant.

- 3) Our results are backed by a novel spectral perturbation analysis that could be a valuable tool for other sparse graph problems.

Organization: In Sections II-A-II-B, we provide more details about SBMs and give a detailed implementation of the spectral clustering. In Section II-C, we give a preliminary polynomial upper bound for the performance of the spectral clustering. In Section II-D, we carry out an oracle analysis to provide intuition on the fundamental limits of the spectral clustering. We establish the upper bound part of Theorem 1 in Section III-A and the lower bound part in Section III-C. The truncated ℓ_2 perturbation analysis of eigenvectors is presented in Section III-B. The proofs of the upper bound and the truncated ℓ_2 perturbation analysis are given in Section IV-A and Section IV-B, respectively. We include the proof of the lower bound in the appendix, along with proofs of the results of Section II and all auxiliary lemmas.

Notation: For any positive integer r , let $[r] := \{1, 2, \dots, r\}$. For any vector x , we denote $\|x\|_\infty := \max_i |x_i|$ to be its ℓ_∞ norm and $\|x\|_1 = \sum_i |x_i|$ to be its ℓ_1 norm. For a matrix A , denote A_i and $A_{\cdot i}$ to be its i th row and column, respectively. We further denote $\|A\|$ to be its operator norm, $\|A\|_F$ to be its Frobenius norm, and $\|A\|_{2,\infty} := \max_i \|A_i\|$ to be its maximum ℓ_2 norms of rows. For any two numbers $a, b \in \mathbb{R}$, we denote $a \wedge b := \min\{a, b\}$ and $a \vee b := \max\{a, b\}$. We denote $\mathbb{I}\{\cdot\}$ as the indicator function. For any two positive integers a, b , we denote I_a to be the $a \times a$ identity matrix and $\mathcal{O}(a, b)$ to be the set of all $a \times b$ matrices with orthogonal columns. For any random vectors X, Y , and Z , we use $X \perp Y|Z$ to mean X and Y are independent conditioned on Z . For any event \mathcal{G} , we denote \mathcal{G}^c to be its complement.

II. PRELIMINARIES

A. Community Detection and Stochastic Block Models

Consider an n -node network with its adjacency matrix denoted by $A \in \{0, 1\}^{n \times n}$ such that $A = A^T$ and $A_{ii} = 0$ for all $i \in [n]$. Under the SBM, all edges $\{A_{ij}\}_{1 \leq i < j \leq n}$ are independent Bernoulli random variables with probabilities depending on the underlying community structure. Let $z^* \in [k]^n$ be a community assignment vector such that each coordinate indicates which community the corresponding node belongs to. We assume

$$\mathbb{E}A_{ij} = \begin{cases} p, & \text{if } z_i^* = z_j^*, \\ q, & \text{o.w.,} \end{cases}$$

for all $1 \leq i < j \leq n$, where $0 < q < p < 1$. That is, nodes are more likely to be connected if they are from the same community: the probability for two nodes to be connected is p if they belong to the same community and is q otherwise. The goal of community detection is to estimate z^* given the network A . Throughout the paper, we assume k , the number of communities, is known.

We denote n_1, \dots, n_k to be community sizes such that $n_a := \sum_{i \in [n]} \mathbb{I}\{z_i^* = a\}$ for all $a \in [k]$. Define $\beta := (\min_{a \in [k]} n_a)/(n/k)$ such that $\beta n/k$ is the smallest community size. For any $z \in [k]^n$, its performance for community

detection can be measured by the following loss function [19]:

$$\ell(z, z^*) := \frac{1}{n} \min_{\phi \in \Phi} \sum_{i \in [n]} \mathbb{I}\{z_i = \phi(z_i^*)\},$$

where $\Phi := \{\phi : \phi \text{ is a bijection from } [k] \text{ to } [k]\}$. It is a value between 0 and 1, giving the proportions of nodes mis-clustered in z compared to z^* .

B. Spectral Clustering

Spectral clustering refers to clustering procedures built upon the eigendecomposition or the singular value decomposition of matrices constructed from data. There exist different variants of spectral clustering for community detection [7], [37], [38], [39], [40]. They differ in the matrix on which the spectral decomposition is applied and in what spectral components are used for the subsequent clustering. The spectral clustering considered in this paper is a popular one and has been widely studied in literature. It contains three steps summarized below with the detailed implementation given in Algorithm 1.

Algorithm 1 Spectral Clustering for Community Detection in Stochastic Block Models

Input: Adjacency matrix $A \in \{0, 1\}^{n \times n}$, number of communities k , threshold τ

Output: Community assignment \hat{z}

- 1 Define $d_i := \sum_{j \neq i} A_{ij}$ for all $i \in [n]$ to be degrees of A . Let \tilde{A} be a trimmed version of A by replacing its i th row and column by 0 whenever $d_i \geq \tau$, for all $i \in [n]$. That is, for all $i, j \in [n]$,

$$\tilde{A}_{ij} := \begin{cases} A_{ij}, & \text{if } d_i, d_j < \tau, \\ 0, & \text{o.w.} \end{cases}$$

- 2 Let the eigendecomposition of \tilde{A} be

$\tilde{A} = \sum_{i \in [n]} \lambda_i u_i u_i^T$ with eigenvalues $\lambda_1 \geq \dots \geq \lambda_n$ and eigenvectors $u_1, \dots, u_n \in \mathbb{R}^n$. Define $U := (u_1, \dots, u_k) \in \mathbb{R}^{n \times k}$ to be the leading eigenspace and $\Lambda := \text{diag}(\lambda_1, \dots, \lambda_k) \in \mathbb{R}^{k \times k}$ to be the diagonal matrix with the leading k eigenvalues. Denote $U_1, \dots, U_n \in \mathbb{R}^{1 \times n}$ to be rows of U .

- 3 Apply k -means on the rows of $U\Lambda \in \mathbb{R}^{n \times k}$ and let \hat{z} be the clustering output. That is,

$$\begin{aligned} &(\hat{z}, \{\hat{\theta}_1, \dots, \hat{\theta}_k\}) \\ &:= \underset{z \in [k]^n, \theta_1, \dots, \theta_k \in \mathbb{R}^{1 \times k}}{\text{argmin}} \sum_{i \in [n]} \|U_i \Lambda - \theta_{z_i}\|^2. \end{aligned} \quad (2)$$

In the first step, we regularize the adjacency matrix by removing high-degree nodes, i.e., nodes with degrees greater or equal to the threshold τ . This step is necessary for sparse networks as the adjacency matrix A is known to be away from its expectation $\mathbb{E}A$ when $p \ll (\log n)/n$. On the contrary, by zeroing out rows and columns of A that correspond to the high-degree nodes, we have a trimmed adjacency matrix \tilde{A} that is highly concentrated around $\mathbb{E}A$ [29], [30]. Primarily

for the sake of theoretical analysis, we choose

$$\tau = 20np,$$

throughout the paper. However, it is conceivable to replace this with a more general form, such as Cnp , where C represents a large constant. It is important to emphasize that the inclusion of p , an unknown parameter, in both $20np$ and Cnp makes the threshold, τ , impractical for direct application. In order to render Algorithm 1 more feasible in practice, one could follow the approach in [19] and [41] by setting τ to be $C'\bar{d}$. Here, \bar{d} stands for the average degree and C' is a sufficiently large constant. Nevertheless, for the purposes of streamlined theoretical exposition, we have opted for $\tau = 20np$ in this paper.

In the second step, we obtain two matrices U, Λ through the eigendecomposition of \tilde{A} , where U is a matrix including the leading k eigenvectors and Λ is a diagonal matrix with the leading k eigenvalues. Throughout the paper, we refer to U as the leading eigenspace of A . Since eigenvectors are not of equal importance, they are weighted with their corresponding eigenvalues and $U\Lambda$ is used for the subsequent clustering.

In the third step, we perform the k -means clustering on rows of $U\Lambda$ which are of dimension k . Compared to the adjacency matrix A , we greatly reduce the dimensionality of data from n to k as usually $k \ll n$ and perform clustering in a low-dimensional space. The k -means clustering returns a partition of the data and cluster centers, which are denoted as \hat{z} and $\{\hat{\theta}_a\}_{a \in [k]}$ respectively.

Algorithm 1 effectively performs a rank- k approximation of \tilde{A} as part of its process. To elaborate, after constructing \tilde{A} in Algorithm 1, applying k -means directly to the rows of $U\Lambda$ is, in fact, equivalent to performing k -means on the rows of the rank- k approximated matrix $U\Lambda U^T \in \mathbb{R}^{n \times n}$. This equivalence arises because the orthonormal columns of U ensure that the Euclidean distance between the rows in $U\Lambda$ remains unchanged in $U\Lambda U^T$. Formally, for any two indices $i, j \in [n]$, the distance property $\|U_i \Lambda - U_j \Lambda\| = \|U_i \Lambda U^T - U_j \Lambda U^T\|$ holds, guaranteeing identical clustering results from k -means on either matrix. Although the two approaches are equivalent in terms of clustering outcome, using $U\Lambda$ instead of $U\Lambda U^T$ for k -means offers computational benefits, as the former reduces the computational load and storage requirements, given the lower dimensionality of $U\Lambda$ compared to $U\Lambda U^T$. As a result, in the third step of Algorithm 1, k -means is applied to $U\Lambda$ instead of the rank- k approximation of \tilde{A} .

The current form of Algorithm 1 is tailored for the standard SBMs introduced in Section II-A. However, it holds potential for adaptation to more complex models such as bipartite SBMs, which are characterized by an asymmetric adjacency matrix with distinct row and column community structures, as studied in [18]. To accommodate the unique features of bipartite graphs, the algorithm requires modification as follows. In the initial step, we compute both row-wise and column-wise degrees to identify and trim high-degree nodes, resulting in a matrix \tilde{A} . Subsequently, instead of eigendecomposition, singular value decomposition (SVD) is applied to \tilde{A} , yielding the left singular matrix U , right singular matrix V , and a diagonal matrix of singular values Λ . This

adaptation leverages the asymmetry of the bipartite structure. Finally, community detection is performed by applying k -means clustering separately to the rows of $U\Lambda$ and the columns of ΛV^T , facilitating the recovery of the distinct row-wise and column-wise community structures.

It should be emphasized that Algorithm 1 is not a novel contribution of this paper. Our primary contribution lies not in the algorithm itself, but in the precise and detailed analysis of its performance.

C. A Polynomial Upper Bound

In this section, we provide preliminary analysis for the spectral clustering and show it achieves a polynomial error rate. We first introduce a matrix $P \in \mathbb{R}^{n \times n}$ defined as

$$P_{ij} := p\mathbb{I}\{z_i^* = z_j^*\} + q\mathbb{I}\{z_i^* \neq z_j^*\}, \forall i, j \in [n].$$

It can be viewed as a population matrix with \tilde{A} being its sample counterpart. Note that P is different from the expectation matrix $\mathbb{E}A$ as the latter has all diagonal entries being zero due to the fact that the network has no self-loops.

The matrix P is rank- k . Let the eigendecomposition of P be $P = \sum_{i \in [k]} \lambda_i^* u_i^* u_i^{*T}$ with eigenvalues $\lambda_1^* \geq \dots \geq \lambda_k^*$ and eigenvectors $u_1^*, \dots, u_k^* \in \mathbb{R}^n$. Define $U^* := (u_1^*, \dots, u_k^*) \in \mathbb{R}^{n \times k}$ to be the leading eigenspace and $\Lambda^* := \text{diag}(\lambda_1^*, \dots, \lambda_k^*) \in \mathbb{R}^{k \times k}$ to be the diagonal matrix with the leading k eigenvalues. Then we have $P = U^* \Lambda^* U^{*T}$ and the matrix $U^* \Lambda^*$ satisfies the following property.

Lemma 1: The matrix $U^* \Lambda^* \in \mathbb{R}^{n \times k}$ has k unique rows. To be more specific, there exist $\theta_1^*, \dots, \theta_k^* \in \mathbb{R}^{1 \times k}$ such that $(U^* \Lambda^*)_i = \theta_{z_i^*}^*$ for all $i \in [n]$. In addition,

$$\|\theta_a^* - \theta_b^*\| = \sqrt{n_a + n_b}(p - q),$$

and $\|\theta_a^*\|^2 = (p^2 - q^2)n_a + q^2n$, for all $a, b \in [k]$ such that $a \neq b$.

Lemma 1 reveals that rows of $U^* \Lambda^*$ are equal if their corresponding nodes belong to the same community and the k unique rows $\{\theta_a^*\}$ are separated from each other. As a result, if the k -means clustering is performed on rows of $U^* \Lambda^*$, we will have a perfect partition of the network with clustering centers $\{\theta_a^*\}$. But this is unrealistic as $U^* \Lambda^*$ is unobserved. A natural idea is to use the eigendecomposition of the adjacency matrix A . Intuitively, if A and P are close, then $U^* \Lambda^*$ is close to its counterpart obtained from A , and consequently the clustering error using the latter is small. However, it is known in random matrix and graph theory that sparse random graphs do not concentrate [29], meaning that A is away from P^* . Recent literature reveals that the removal of the high degree vertices enforces concentration [29], which motivates the use of \tilde{A} instead of A in Algorithm 1. The concentration of \tilde{A} around P (note that P and $\mathbb{E}A$ are nearly equal as they only differ in diagonal entries) is given in the following Lemma 2.

Lemma 2: There exists a constant $C_0 > 0$ such that

$$\|\tilde{A} - \mathbb{E}A\| \leq C_0 \sqrt{np} \quad (3)$$

with probability at least $1 - 2n^{-3}$.

In Lemma 2, C_0 represents an unspecified absolute constant. Lemma 2 is given as Lemma 12 and proved in [30]. Despite that [30] only states that the upper bound holds with probability $1 - o(1)$, its proof gives an explicit expression for the probability that is at least $1 - 2n^{-3}$. However, [30] does not give an explicit value for C_0 . By scrutinizing its proof, $1 - 2n^{-3}$ in Lemma 2 can be generalized to $1 - n^{-r}$ for any constant $r > 0$, and then C_0 can be denoted as $C_0(r)$, a function of r . A similar result is given as Theorem 1.1 and proved in [29]. The proofs in [29] and [30] are lengthy and technical. For these reasons, and to maintain focus on the core contributions of our work, we choose not to include the proof of the lemma in this paper and refer readers to these sources for a broader context of graph concentration.

In the remaining part of the paper, we will analyze the performance of the spectral clustering under the with-high-probability event that \tilde{A} is well-concentrated around $\mathbb{E}A$. Denote an event

$$\mathcal{F} := \mathbb{I}\{\|\tilde{A} - \mathbb{E}A\| \leq C_0\sqrt{np}\}, \quad (4)$$

where C_0 is the constant from the statement of Lemma 2. Then $\mathbb{P}(\mathcal{F}) \geq 1 - 2n^{-3}$. Under \mathcal{F} , by classical spectral perturbation theory, we immediately have the following preliminary result.

Proposition 1: Assume the event \mathcal{F} holds. There exist constants $C_1, C_2 > 0$ and some $\phi \in \Phi$ such that if $\frac{n(p-q)^2}{\beta^{-2}k^3p} \geq C_1$, we have

$$\frac{1}{n} \sum_{i \in [n]} \mathbb{I}\{\hat{z}_i \neq \phi(z_i^*)\} \leq C_2 \frac{k^2 \beta^{-1} p}{n(p-q)^2}, \quad (5)$$

$$\text{and } \max_{a \in [k]} \left\| \hat{\theta}_{\phi(a)} U^T - \theta_a^* U^{*T} \right\| \leq C_2 \beta^{-0.5} k \sqrt{p}. \quad (6)$$

The upper bound in (5) is essentially equal to the reciprocal of $n(p-q)^2/p$, a quantity regarded as the signal-to-noise ratio in the community detection literature [26]. Hence, (5) decreases polynomially as $n(p-q)^2/p$ grows and we refer to it as a polynomial upper bound. Proposition 1 also gives an upper bound for the deviation between cluster centers $\{\hat{\theta}_a\}$ and their population counterpart $\{\theta_a^*\}$. Results similar to Proposition 1 for spectral clustering can be found in [16], [31], [33]. Proposition 1 serves as the starting point for our further analysis toward Theorem 1.

D. Oracle Analysis and Exponents

In this section, we provide heuristic arguments to explain why the spectral clustering has exponential error bounds and to derive the exponent J_{\min} . For any two positive integers m_1, m_2 , define

$$J_{m_1, m_2, p, q} := \max_{t > 0} \left((m_1 - m_2) t \frac{p+q}{2} - m_1 \log(qe^t + 1 - q) - m_2 \log(pe^{-t} + 1 - p) \right). \quad (7)$$

Then $J_{\min} = \min_{1 \leq a \neq b \leq k} J_{n_a, n_b, p, q}$. For simplicity, we consider a two-community SBM. In addition, instead of analyzing \hat{z} , we study a simplified procedure $\tilde{z} \in \{1, 2\}^n$ defined as

follows:

$$\tilde{z}_i := \begin{cases} 1, & \text{if } \|A_i U^* - \theta_1^*\| \leq \|A_i U^* - \theta_2^*\|, \\ 2, & \text{o.w.} \end{cases}$$

We refer to \tilde{z} as an oracle estimator since it involves the unknown U^* and θ_1^*, θ_2^* and is not practical. Nevertheless, \tilde{z} is closely related to \hat{z} as we elaborate below. Note that \hat{z} performs clustering on rows of $\tilde{A}U$ as $\tilde{A}U = U\Lambda$. The k -means clustering (2) implies \hat{z}_i is equal to 1 if $\|\tilde{A}_i U - \hat{\theta}_1\| \leq \|\tilde{A}_i U - \hat{\theta}_2\|$ and is equal to 2 otherwise for each $i \in [n]$. As a comparison, \tilde{z} has a similar form but with $\tilde{A}U$ replaced by AU^* and $\hat{\theta}_1, \hat{\theta}_2$ replaced by θ_1^*, θ_2^* . Since θ_1^*, θ_2^* are provided, \tilde{z} is more of a classification procedure than a clustering method. Despite all these discrepancies, \tilde{z} captures the key ingredient of \hat{z} and analyzing \tilde{z} reveals fundamentally important properties of the spectral clustering.

The following proposition characterizes the statistical accuracy of the oracle estimator with both upper and lower bounds. From Proposition 2, the oracle estimator has an exponential classification error with J_{\min} being the exact asymptotic exponent. Though it only considers $k = 2$ case, it can be generalized to multi-community cases with J_{\min} appearing in the exponent.

Proposition 2: Consider a two-community SBM with community sizes n_1, n_2 both of the order n . In addition, assume $0 < q < p \leq 1/2$, p, q are of the same order, and $\frac{n(p-q)^2}{p} \rightarrow \infty$, we have

$$\mathbb{E} \left(\frac{1}{n} \sum_{i \in [n]} \mathbb{I}\{\tilde{z}_i \neq z_i^*\} \right) \leq \exp(-(1 - o(1))(J_{n_1, n_2, p, q} \wedge J_{n_2, n_1, p, q})),$$

and

$$\mathbb{E} \left(\frac{1}{n} \sum_{i \in [n]} \mathbb{I}\{\tilde{z}_i \neq z_i^*\} \right) \geq \exp(-(1 + o(1))(J_{n_1, n_2, p, q} \wedge J_{n_2, n_1, p, q})).$$

To explain why $J_{n_1, n_2, p, q} \wedge J_{n_2, n_1, p, q}$ appears in Proposition 2, let us consider any $i \in [n]$ such that $z_i^* = 2$. By simple algebra (see proof of Proposition 2), the event $\mathbb{I}\{\tilde{z}_i \neq z_i^*\}$ can be written equivalently as

$$\mathbb{I} \left\{ \sum_{j: z_j^* = 1} A_{ij} - \sum_{j \neq i: z_j^* = 2} A_{ij} \geq (n_1 - n_2) \frac{p+q}{2} \right\}. \quad (8)$$

Note that for any $j \in [n]$, A_{ij} is a Bernoulli random variable with probability q if $z_j^* = 1$ and p if $z_j^* = 2$ and $j \neq i$. In addition, $|\{j : z_j^* = 1\}| = n_1$ and $|\{j \neq i : z_j^* = 2\}| = n_2 - 1$. Let $\{X_i\}$ and $\{Y_j\}$ be independent Bernoulli random variables with probabilities q, p respectively. Then $\mathbb{E} \mathbb{I}\{\tilde{z}_i \neq z_i^*\}$ is essentially equal to

$$\mathbb{P} \left(\sum_{l \in [n_1]} X_l - \sum_{j \in [n_2]} Y_j \geq (n_1 - n_2) \frac{p+q}{2} \right),$$

where we ignore a minor difference between n_2 and $n_2 - 1$ which is negligible. For this tail probability, a direct application of Chernoff bound leads to an upper bound $\exp(-J_{n_1, n_2, p, q})$, and its lower bound $\exp(-(1 + o(1))J_{n_1, n_2, p, q})$ can be established using the Cramér-Chernoff argument. Similarly, $J_{n_2, n_1, p, q}$ appears in the analysis for any $i \in [n]$ such that $z_i^* = 1$. Between these two exponents, the smaller one of $J_{n_1, n_2, p, q}$ and $J_{n_2, n_1, p, q}$ dominates and leads to Proposition 2.

III. MAIN RESULTS

A. Upper Bound

In this section, we present one main result of the paper: a sharp upper bound for the performance of the spectral clustering. Theorem 2 is essentially the upper bound part of Theorem 1 but with an explicit formula for the $o(1)$ term in the exponent.

Theorem 2: Under the assumption that β^{-1} , $k = O(1)$, $0 < q < p \leq 1/10$, $\frac{p}{q} = O(1)$, and $\frac{n(p-q)^2}{p} \rightarrow \infty$, there exists some constant $C > 0$ such that

$$\mathbb{E} \ell(\hat{z}, z^*) \leq \exp \left(- \left(1 - C \left(\log \left(\frac{n(p-q)^2}{p} \right) \right)^{-\frac{1}{2}} \right) J_{\min} \right) + 2n^{-3}.$$

The upper bound in Theorem 2 involves two terms, the exponential term is essentially from $\mathbb{E} \ell(\hat{z}, z^*) \mathbb{I}\{\mathcal{F}\}$ where we study $\ell(\hat{z}, z^*)$ under the event \mathcal{F} and the term $2n^{-3}$ comes from $\mathbb{P}(\mathcal{F}^c) \leq 2n^{-3}$ (see Lemma 2). Since $2n^{-3}$ can be improved into n^{-C} for any constant $C > 0$ by scrutinizing the proof of Lemma 2, it should be generally ignored. By Lemma 13, J_{\min} can be shown to be of the order $n(p-q)^2/p$. Hence, $n(p-q)^2/p \rightarrow \infty$ is the sufficient condition for \hat{z} to have a vanishing error. When $J_{\min} \geq (1 + \epsilon) \log n$ for any constant $\epsilon > 0$, Theorem 2 immediately implies that \hat{z} achieves the exact recovery with high probability. Theorem 2 requires that the number of communities k remains constant. However, our preliminary result in Proposition 1, allows k to increase with n , provided it does not grow faster than $(\frac{n(p-q)^2}{\beta^{-2}p})^{\frac{1}{3}}$. This more restrictive condition on k in our main results arises from the limitations of our current proof techniques.

The proof of Theorem 2 is quite involved. Below we give an overview of the proof and highlight challenges we face and new ideas and techniques we use to address them. Since $\tilde{A}U = U\Lambda$, the k -means clustering is performed on rows of $\tilde{A}U$ and \hat{z}_i satisfies $\hat{z}_i = \arg\min_{a \in [k]} \|\tilde{A}_i \cdot U - \hat{\theta}_a\|$ for each $i \in [n]$. Consider any node $i \in [n]$ and any $a \neq z_i^*$. The node will be incorrectly clustered if $\|\tilde{A}_i \cdot U - \hat{\theta}_a\| \leq \|\tilde{A}_i \cdot U - \hat{\theta}_{z_i^*}\|$ happens, an event that is about $(\tilde{A}_i - \mathbb{E}A_i)U$. To analyze it, we further decompose it into two terms: one is essentially in a form of $(A_i - \mathbb{E}A_i)U^*$ and the other one is $(\tilde{A}_i - \mathbb{E}A_i)(I - U^*U^{*T})U$.

The analysis for the term $(A_i - \mathbb{E}A_i)U^*$ is relatively straightforward. It eventually leads to the following event

$$\mathbb{I} \left\{ \sum_{j: z_j^* = a} A_{ij} - \sum_{j: z_j^* = z_i^*, j \neq i} A_{ij} \geq (n_a - n_{z_i^*}) \frac{p+q}{2} - o(1)(n_a + n_{z_i^*}) \frac{p-q}{2} \right\}, \quad (9)$$

which mimics the event (8) of the oracle estimator. The existence of $o(1)(n_a + n_{z_i^*})(p-q)/2$ is due to the discrepancy between the spectral clustering and the oracle estimator. The probability of the event (9) leads to an upper bound $\exp(-(1 - o(1))J_{n_a, n_{z_i^*}, p, q})$. Going through all $i \in [n]$ and $a \in [k]$, the largest one is $\exp(-(1 - o(1))J_{\min})$ and appears in the upper bound of Theorem 2.

The proof of Theorem 2 mostly focuses on analyzing tail probabilities of $(\tilde{A}_i - \mathbb{E}A_i)(I - U^*U^{*T})U$ which is the most difficult and technical part toward establishing Theorem 2. There are two main challenges. The first challenge is that \tilde{A}_i and U are heavily dependent on each other. This challenge can be addressed by a refined leave-one-out technique to decouple the dependence. The second challenge is more critical and is the reason why we need to develop new techniques. For the purpose of illustration, let us consider a simplified setting: we want to study tail probabilities of $(X - \mathbb{E}X)u$ where $X \in \{0, 1\}^{1 \times n}$ with $X_i \stackrel{iid}{\sim} \text{Bernoulli}(p)$ and is independent of u , a column of $(I - U^*U^{*T})U$. Existing literature [24], [34], [35] typically applies classical tail probabilities of Bernoulli random variables (e.g., Bernstein inequality) which involve both $\|u\|$ and $\|u\|_\infty$. Between these two norms, the former one is well-controlled but the latter one is essentially $\|U\|_{2, \infty}$ that has some inevitable $\log n$ factor. To deal with this $\log n$ factor and to derive meaningful tail probabilities, p has to be much bigger than $(\log n)/n$. However, in this paper, we consider sparse networks with the connectivity probability allowed to be far smaller than $(\log n)/n$ and the aforementioned analysis breaks down.

To analyze $(\tilde{A}_i - \mathbb{E}A_i)(I - U^*U^{*T})U$, we develop a new technical tool that avoids the use of $\|u\|_\infty$ by truncating its coordinates. More accurately, to avoid dealing with $\|U\|_{2, \infty}$, we truncate rows of U if their ℓ_2 norms are above a certain threshold. Let $t > 0$ and define a mapping $f_t : \mathbb{R}^{1 \times k} \rightarrow \mathbb{R}^{1 \times k}$ such that

$$f_t(x) = \begin{cases} x, & \text{if } \|x\| \leq t, \\ tx / \|x\|, & \text{if } \|x\| > t. \end{cases} \quad (10)$$

That is, $f_t(x)$ truncates x if $\|x\| > t$ so that $\|f_t(x)\| \leq t$ is always satisfied. If $k = 1$, f_t is a truncation operator such that $f_t(x) = x \mathbb{I}\{|x| \leq t\} + t \mathbb{I}\{x > t\} - t \mathbb{I}\{x < -t\}$ for any scalar x . Apply f_t to rows of U and we get a matrix $\bar{U} \in \mathbb{R}^{n \times k}$ defined as

$$\bar{U}_i := f_t(U_i), \forall i \in [n]. \quad (11)$$

The definition of \bar{U} ensures $\|\bar{U}\|_{2, \infty} \leq t$. Note that \bar{U} depends on the value of t . In the proof of Theorem 2, we set $t =$

t_0 where t_0 is defined later in (13). With \bar{U} , we decompose $(\tilde{A}_i - \mathbb{E}A_i)(I - U^*U^{*T})U$ into

$$(\tilde{A}_i - \mathbb{E}A_i)(I - U^*U^{*T})U = (\tilde{A}_i - \mathbb{E}A_i)(I - U^*U^{*T})\bar{U} + (\tilde{A}_i - \mathbb{E}A_i)(I - U^*U^{*T})(U - \bar{U}). \quad (12)$$

We now use classical concentration inequalities to analyze $(\tilde{A}_i - \mathbb{E}A_i)(I - U^*U^{*T})\bar{U}$ (after decoupling the dependence between \tilde{A}_i and \bar{U} by the leave-one-out technique) thanks to that the controlled $\|\bar{U}\|_{2,\infty}$. It leads to negligible tail probabilities that can be absorbed into $\exp(-(1-o(1))J_{\min})$. Aggregating all $i \in [n]$, the impact of $(\tilde{A}_i - \mathbb{E}A_i)(I - U^*U^{*T})(U - \bar{U})$, the second term in (12), turns out to be related to a truncated ℓ_2 norm $\sum_{i \in [n]} \|U_i\|^2 \mathbb{I}\{\|U_i\| \geq t_0\}$. According to Theorem 3 below, this truncated ℓ_2 norm is exponentially small. With the help of Theorem 3, the contribution of $\{(\tilde{A}_i - \mathbb{E}A_i)(I - U^*U^{*T})(U - \bar{U})\}$ toward $\ell(\hat{z}, z^*)$ is also negligible and can be absorbed into $\exp(-(1-o(1))J_{\min})$ as well. In this way, we derive the exponential term $\exp(-(1-o(1))J_{\min})$ in Theorem 2.

B. Truncated ℓ_2 Perturbation Analysis for Eigenspaces

Theorem 3: Define

$$t_0 := \frac{160k^2}{\beta} \sqrt{\frac{k}{\beta n}}. \quad (13)$$

Under the assumption that $\beta^{-1}, k = O(1), 0 < q < p \leq 1/2$, and $\frac{n(p-q)^2}{p} \rightarrow \infty$, we have

$$\mathbb{E} \left(\sum_{i \in [n]} \|U_i\|^2 \mathbb{I}\{\|U_i\| \geq t_0\} \right) \mathbb{I}\{\mathcal{F}\} \leq t_0^2 n \exp \left(-\frac{3n(p-q)^2}{2p} \right).$$

Theorem 3 provides a truncated ℓ_2 norm for U and is crucial to establishing Theorem 2. The choice of t_0 in (13) is carefully chosen. Note that $\|U^*\|_{2,\infty} = \sqrt{k/(\beta n)}$ according to Lemma 3. When β, k are both constants, t_0 is equal to $\|U^*\|_{2,\infty}$ multiplied by a large constant. Since U can be seen as a perturbation of U^* , the truncated ℓ_2 norm can also be interpreted as a perturbation bound for U and U^* . Particularly, it focuses on rows of U with norms that far exceed the baseline $\|U^*\|_{2,\infty}$. Theorem 3 immediately implies the number of such rows is exponentially small.

Corollary 1: Under the same assumption as in Theorem 3, we have

$$\mathbb{E} \left(\sum_{i \in [n]} \mathbb{I}\{\|U_i\| \geq t_0\} \right) \mathbb{I}\{\mathcal{F}\} \leq n \exp \left(-\frac{3n(p-q)^2}{2p} \right).$$

Theorem 3 and Corollary 1 together provide insights on why the eigenvector truncation idea works in analyzing $(\tilde{A}_i - \mathbb{E}A_i)(I - U^*U^{*T})U$ in the proof of Theorem 2. Let us revisit the simplified setting $(X - \mathbb{E}X)u$ discussed above. The magnitude of $\|u\|_\infty$ does break down classical concentration inequalities. However, these concentration inequalities are usually sharp for the worst case: their upper bounds hold for all

weight vectors u with given $\|u\|$ and $\|u\|_\infty$. On the contrary, the particular u appearing in our analysis is very special: it has a small number of high magnitude coordinates (Corollary 1) which have an exponentially small aggregated ℓ_2 norm (Theorem 3), an important property completely ignored if classical concentration inequalities are applied. Instead, in our analysis, we fully exploit this structure by singling out high-magnitude coordinates of u using the eigenvector truncation and eliminate the use of $\|u\|_\infty$. This is the intuition behind (12) which is crucial to proving Theorem 2. The idea of eigenvector truncation, the decomposition (12), and Theorem 3 might be useful for spectral perturbation analysis of other random matrix problems and models.

The proof of Theorem 3 is also complicated. Since $U = \tilde{A}U\Lambda^{-1}$, for each $i \in [n]$, the event $\|U_i\|^2 \mathbb{I}\{\|U_i\| \geq t_0\}$ is essentially also about $(\tilde{A}_i - \mathbb{E}A_i)U$, the same as in the proof of Theorem 2. As a result, we face the same challenges that appear in the analysis of Theorem 2, and we address them with the same eigenvector truncation arguments and techniques. Similar to how we establish Theorem 2, we decompose $(\tilde{A}_i - \mathbb{E}A_i)U$ into two parts: one eventually leads to a well-controlled tail probability of $(\tilde{A}_i - \mathbb{E}A_i)U^*$, and the other one involves $(\tilde{A}_i - \mathbb{E}A_i)(I - U^*U^{*T})U$. For the latter one, we use (12) again, handle the quantity $(\tilde{A}_i - \mathbb{E}A_i)(I - U^*U^{*T})\bar{U}$ with the help of the leave-one-out technique, and aggregate $(\tilde{A}_i - \mathbb{E}A_i)(I - U^*U^{*T})\bar{U}$ across all $i \in [n]$ which leads to $\sum_{i \in [n]} \|U_i\|^2 \mathbb{I}\{\|U_i\| \geq t_0\}$. The remaining part of the proof is different from that of Theorem 2. So far, we obtain an inequality: its left-hand side is our target of Theorem 3, the truncated ℓ_2 norm $\sum_{i \in [n]} \|U_i\|^2 \mathbb{I}\{\|U_i\| \geq t_0\}$; its right-hand side involves various terms with one being the truncated ℓ_2 norm as well. The one appearing on the right-hand side (from the aggregation of $\{(\tilde{A}_i - \mathbb{E}A_i)(I - U^*U^{*T})\bar{U}\}$) can be shown to have a small constant factor and consequently can be absorbed into the one on the left-hand side. The inequality then becomes an upper bound for the truncated ℓ_2 norm and leads to Theorem 3.

C. Lower Bound

Following the proof of Theorem 2 with some modifications, we are able to obtain a matching lower bound presented below in Theorem 4. It corresponds to the lower bound part of Theorem 1.

Theorem 4: Under the same assumption as in Theorem 2, there exists some constant $C' > 0$ such that

$$\mathbb{E}\ell(\hat{z}, z^*) \geq \exp \left(- \left(1 + C' \left(\log \left(\frac{n(p-q)^2}{p} \right) \right)^{-\frac{1}{4}} \right) J_{\min} \right) - 2n^{-3}.$$

By Lemma 13, J_{\min} is of the order $n(p-q)^2/p$. Theorem 4 indicates $n(p-q)^2/p \rightarrow \infty$ is the necessary condition for \hat{z} to have a vanishing error. When $J_{\min} \leq (1-\epsilon) \log n$ for any constant $\epsilon > 0$, Theorem 2 immediately implies that \hat{z} has a constant probability of not achieving the exact recovery.

Theorem 4 complements Theorem 2 and together they provide a precise characterization of the performance of the

spectral clustering \hat{z} , which is referred to as its fundamental limit in this paper, demonstrating its ability and limit. This differs from the information-theoretical perspective on SBMs, referred to as fundamental limits of SBMs in [1] and minimax rates of SBMs in [26]. Specifically, [1] establishes a sharp threshold for exact community recovery, delineating a boundary where below it, no algorithm can succeed, and above it, there exists an algorithm capable of exact recovery. Reference [26] gives the minimax risk of the community detection problem: $\inf_{\hat{z}} \sup_{z^*} \mathbb{E} \ell(\hat{z}, z^*)$ where the infimum is over all possible algorithms. In contrast, our focus is on the spectral clustering algorithm \hat{z} , with Theorems 2 and 4 providing bounds for $\mathbb{E} \ell(\hat{z}, z^*)$.

From Theorems 2 and 4, the efficacy of \hat{z} for exact recovery hinges on whether $J_{\min}/\log n > 1 + \epsilon$ or $J_{\min}/\log n < 1 - \epsilon$, for some constant $\epsilon > 0$. For SBMs where all connectivity probabilities are of order $(\log n)/n$, [1] showed that exact recovery is feasible by some algorithm if and only if the Chernoff-Hellinger divergence $D_+ > 1$. Although both $J_{\min}/\log n$ and D_+ relate to tail probabilities of Bernoulli random variables, they are distinct; the former is specific to the performance of spectral clustering, while the latter arises from the information-theoretic analysis.

IV. PROOFS OF THEOREM 2 AND THEOREM 3

In this section, we give detailed proofs of Theorem 2 and Theorem 3. Before giving the proofs, we first introduce some useful concepts and tools and state some important properties.

Preliminary Results From Proposition 1: Recall the definition of the event \mathcal{F} in (4). Under this event, Proposition 1 gives a preliminary analysis for the performance of the spectral clustering, showing that there exists some constant $C_2 > 0$ and some $\phi \in \Phi$ such that (5) and (6) hold as $\frac{n(p-q)^2}{\beta^{-2}k^3p} \rightarrow \infty$. Without loss of generality, we can assume $\phi = \text{Id}$. Then (6) leads to

$$\max_{a \in [k]} \left\| \hat{\theta}_a U^T - \theta_a^* U^{*T} \right\| \leq C_2 \beta^{-0.5} k \sqrt{p}. \quad (14)$$

In addition, we have

$$\mathbb{E} \ell(\hat{z}, z^*) \mathbb{I}\{\mathcal{F}\} = \mathbb{E} \frac{1}{n} \sum_{i \in [n]} \mathbb{I}\{\hat{z}_i \neq z_i^*\} \mathbb{I}\{\mathcal{F}\}. \quad (15)$$

Facts About U : Define

$$S := \{i \in [n] : d_i \geq \tau\},$$

to be the set that includes all indexes of rows and columns zeroed out in the spectral clustering where the value of τ is given in Algorithm 1. Then \tilde{A} can be written equivalently as

$$\tilde{A}_{ij} = A_{ij} \mathbb{I}\{i \notin S \text{ and } j \notin S\}, \forall i, j \in [n].$$

Since U is the leading eigenspace of \tilde{A} , we have

$$U_i = 0, \forall i \in S, \quad (16)$$

and consequently $\|U_i\| = 0, \forall i \in S$. In addition, we have

$$(\tilde{A}_i - \mathbb{E} A_i)U = (A_i - \mathbb{E} A_i)U, \forall i \notin S. \quad (17)$$

This is because due to (16), the left-hand side of (17) is equal to $\sum_{j \notin S} (\tilde{A}_{ij} - \mathbb{E} A_{ij})U_j = \sum_{j \notin S} (A_{ij} - \mathbb{E} A_{ij})U_j$, which is equal to the right-hand side of (17).

Row-Wise Truncation of U : In the proofs of the main theorems, we need to study quantities that involve rows of A and U . To avoid dealing with $\|U\|_{2,\infty}$, we truncate rows of U if their ℓ_2 norms are above a certain threshold. Recall the definitions of f_t and \bar{U} in (10) and (11). Since $\|f_t(x) - x\| \leq \|x\| \mathbb{I}\{\|x\| \geq t\}$, we have

$$\|U - \bar{U}\|_F^2 = \sum_{i \in [n]} \|U_i - \bar{U}_i\|^2 \leq \sum_{i \in [n]} \|U_i\|^2 \mathbb{I}\{\|U_i\| \geq t\}. \quad (18)$$

Note that \bar{U} depends on the value of t . In the proofs of the main theorems, we will first consider any $t > 0$ and then set $t = t_0$.

Leave-One-Out Counterparts of U and \bar{U} : One challenge is that rows of \tilde{A} and U are not independent of each other. To decouple the dependence, we introduce leave-one-out counterparts of U and \bar{U} . For any $i \in [n]$, define $\tilde{A}^{(i)} \in \{0, 1\}^{n \times n}$ to be the leave-one-out counterpart of \tilde{A} such that

$$\tilde{A}_{jl}^{(i)} := \tilde{A}_{jl} \mathbb{I}\{j \neq i \text{ and } l \neq i\}, \forall j, l \in [n]. \quad (19)$$

That is, $\tilde{A}^{(i)}$ is obtained from \tilde{A} by zeroing out its i th row and column. Let $U^{(i)} \in \mathbb{R}^{n \times k}$ be the leading eigenspace of $\tilde{A}^{(i)}$. Under the event \mathcal{F} and $\frac{n(p-q)^2}{\beta^{-2}k^4p} \rightarrow \infty$, conditions in Lemma 7 are satisfied, which leads to

$$\|U - U^{(i)} W^{(i)}\|_F \leq 6k^{1.5} \|U_i\|, \quad (20)$$

for some orthogonal matrix $W^{(i)} \in \mathcal{O}(k, k)$. Apply f_t to rows of $U^{(i)}$ and we have $\bar{U}^{(i)} \in \mathbb{R}^{n \times k}$:

$$\bar{U}_{j \cdot}^{(i)} := f_t(U_{j \cdot}^{(i)}), \forall j \in [n]. \quad (21)$$

Since $W^{(i)}$ is an orthogonal matrix, for any $j \in [n]$, we have $\|U_{j \cdot}^{(i)} W^{(i)}\| = \|U_{j \cdot}^{(i)}\|$ and consequently, $\bar{U}_{j \cdot}^{(i)} W^{(i)} = f_t(U_{j \cdot}^{(i)} W^{(i)})$. By Lemma 8, we have $\|\bar{U}_{j \cdot} - \bar{U}_{j \cdot}^{(i)} W^{(i)}\| = \|f_t(U_{j \cdot}) - f_t(U_{j \cdot}^{(i)} W^{(i)})\| \leq \|U_{j \cdot} - U_{j \cdot}^{(i)} W^{(i)}\|$. Then (20) gives

$$\|\bar{U} - \bar{U}^{(i)} W^{(i)}\|_F \leq \|U - U^{(i)} W^{(i)}\|_F \leq 6k^{1.5} \|U_i\|. \quad (22)$$

Similar to \bar{U} , $\bar{U}^{(i)}$ depends on the value of t . In the proofs of the main theorems, we will first consider any $t > 0$ and then set $t = t_0$.

Decomposition of Rows of \tilde{A} : For each $i \in [n]$, \tilde{A}_i and the leave-one-out eigenspace $\bar{U}^{(i)}$ still depend on each other, mainly due to the removal of high-degree nodes and S . To further decouple the dependence, we need to have a decomposition for \tilde{A}_i . Define a set

$$S_i := \left\{ j \in [n] : j \neq i \text{ and } \sum_{l \neq i, j} A_{jl} \geq \tau - 1 \right\}. \quad (23)$$

By its definition, S_i is independent of A_i . We also have

$$S \cup \{i\} \subset S_i \cup \{i\}. \quad (24)$$

This is because for each $j \notin i$, we have $d_j - (\sum_{l \neq i, j} A_{jl}) = A_{ji} = A_{ij} \in \{0, 1\}$. As a result, if $j \in S$, it must be in S_i as well; if $j \in S_i$, it is in S if and only if $A_{ij} = 1$. This leads to the independence between $S \cup \{i\}$ and coordinates in A_i , with indexes not in $S_i \cup \{i\}$, conditioned on the set $S_i \cup \{i\}$. Define \tilde{A}_i, \tilde{P}_i such that

$$\begin{aligned} \tilde{A}_{ij} &:= A_{ij} \mathbb{I}\{j \neq i \text{ and } j \in S_i^c\} \\ \text{and } \tilde{P}_{ij} &:= P_{ij} \mathbb{I}\{j \neq i \text{ and } j \in S_i^c\}, \forall j \in [n]. \end{aligned} \quad (25)$$

Then the aforementioned independence can be stated equivalently as

$$(\tilde{A}_i - \tilde{P}_i) \perp S \cup \{i\} | S_i \cup \{i\}.$$

By its definition in (19), $\tilde{A}^{(i)}$ is obtained from A by zeroing out rows and columns in $S \cup \{i\}$. In addition, $\bar{U}^{(i)}$ is a function of $\tilde{A}^{(i)}$. Hence,

$$(\tilde{A}_i - \tilde{P}_i) \perp \tilde{A}^{(i)} | S_i \cup \{i\} \text{ and } (\tilde{A}_i - \tilde{P}_i) \perp \bar{U}^{(i)} | S_i \cup \{i\}, \quad (26)$$

for all $i \in [n]$.

Decouple Dependence Between Rows of \tilde{A} and \mathcal{F} : For each $i \in [n]$, $\tilde{A}_i - \tilde{P}_i$ introduced above and the event \mathcal{F} are dependent on each other. To decouple this dependence, we introduce

$$\mathcal{F}_i := \mathbb{I}\{\|\tilde{A}^{(i)} - \mathbb{E}A\| \leq (C_0 + 2)\sqrt{np}\}. \quad (27)$$

Denote $[\mathbb{E}A]^{(i)} \in [0, 1]^{n \times n}$ to be a matrix that is equal to $\mathbb{E}A$ but with its i th row and column zeroed out. Then $\tilde{A}^{(i)} - [\mathbb{E}A]^{(i)}$ is equal to $\tilde{A} - \mathbb{E}A$ but with its i th row and column zeroed out. Then we have $\|\tilde{A}^{(i)} - [\mathbb{E}A]^{(i)}\| \leq \|\tilde{A} - \mathbb{E}A\|$ and consequently $\|\tilde{A}^{(i)} - \mathbb{E}A\| \leq \|\tilde{A} - \mathbb{E}A\| + \|\mathbb{E}A - [\mathbb{E}A]^{(i)}\| \leq \|\tilde{A} - \mathbb{E}A\| + 2\|\mathbb{E}A_i\| \leq \|\tilde{A} - \mathbb{E}A\| + 2\sqrt{np}$. Hence

$$\mathbb{I}\{\mathcal{F}\} \leq \mathbb{I}\{\mathcal{F}_i\}, \forall i \in [n].$$

Since \mathcal{F}_i is a function of $\tilde{A}^{(i)}$, from (26) we have

$$(\tilde{A}_i - \tilde{P}_i) \perp \mathcal{F}_i | S_i \cup \{i\}, \quad (28)$$

for all $i \in [n]$. By Lemma 9, we have

$$\left\| (I_n - U^* U^{*T}) \bar{U}^{(i)} \right\|_{2, \infty} \leq t + \sqrt{\frac{k}{\beta n}} \quad (29)$$

for all $i \in [n]$. In addition, under the event \mathcal{F}_i , by Lemma 9, as $\frac{n(p-q)^2}{\beta^2 k^2 p} \rightarrow \infty$, we have

$$\left\| (I_n - U^* U^{*T}) \bar{U}^{(i)} \right\|_F \leq \frac{2\sqrt{2}(C_0 + 3)k\sqrt{kn}p}{\beta n(p-q)} \quad (30)$$

for all $i \in [n]$ when $t > \sqrt{k/(\beta n)}$.

A. Proof of Theorem 2

Note that

$$\begin{aligned} & \mathbb{E} \ell(\hat{z}, z^*) \\ & \leq \mathbb{E} \ell(\hat{z}, z^*) \mathbb{I}\{\mathcal{F}\} + \mathbb{P}(\mathcal{F}^c) \\ & \leq \mathbb{E} \frac{1}{n} \sum_{i \in [n]} \mathbb{I}\{\hat{z}_i \neq z_i^*\} \mathbb{I}\{\mathcal{F}\} + 2n^{-3} \end{aligned}$$

$$\begin{aligned} & = \mathbb{E} \frac{1}{n} \sum_{i \notin S} \mathbb{I}\{\hat{z}_i \neq z_i^*\} \mathbb{I}\{\mathcal{F}\} + \frac{1}{n} \mathbb{E}|S| + 2n^{-3} \\ & \leq \frac{1}{n} \mathbb{E} \sum_{i \notin S} \mathbb{I}\{\hat{z}_i \neq z_i^*\} \mathbb{I}\{\mathcal{F}\} + \exp(-128np) + 2n^{-3}, \quad (31) \end{aligned}$$

where the second inequality is by (15) and the last inequality is by Lemma 6. In the following proof, we are going to establish an upper bound for $\mathbb{E} \sum_{i \notin S} \mathbb{I}\{\hat{z}_i \neq z_i^*\} \mathbb{I}\{\mathcal{F}\}$.

Now consider a fixed $i \notin S$. According to the objective function of the k -means, we have the following inequality

$$\begin{aligned} & \mathbb{I}\{\hat{z}_i \neq z_i^*\} \mathbb{I}\{\mathcal{F}\} \\ & \leq \mathbb{I}\left\{ \min_{a \neq z_i^*} \|U_i \Lambda - \hat{\theta}_a\| \leq \|U_i \Lambda - \hat{\theta}_{z_i^*}\| \right\} \mathbb{I}\{\mathcal{F}\} \\ & \leq \sum_{a \neq z_i^*} \mathbb{I}\left\{ \|U_i \Lambda - \hat{\theta}_a\| \leq \|U_i \Lambda - \hat{\theta}_{z_i^*}\| \right\} \mathbb{I}\{\mathcal{F}\}. \quad (32) \end{aligned}$$

Further consider a fixed $a \in [k] \setminus \{z_i^*\}$. We are going to study the event that $U_i \Lambda$ is closer to an incorrect center $\hat{\theta}_a$ than the correct center $\hat{\theta}_{z_i^*}$. We have

$$\begin{aligned} & \mathbb{I}\left\{ \|U_i \Lambda - \hat{\theta}_a\| \leq \|U_i \Lambda - \hat{\theta}_{z_i^*}\| \right\} \mathbb{I}\{\mathcal{F}\} \\ & = \mathbb{I}\left\{ \langle U_i \Lambda - \hat{\theta}_{z_i^*}, \hat{\theta}_a - \hat{\theta}_{z_i^*} \rangle \geq \frac{1}{2} \|\hat{\theta}_a - \hat{\theta}_{z_i^*}\|^2 \right\} \mathbb{I}\{\mathcal{F}\}. \quad (33) \end{aligned}$$

We are going to have a decomposition of $\langle U_i \Lambda - \hat{\theta}_{z_i^*}, \hat{\theta}_a - \hat{\theta}_{z_i^*} \rangle$. Due to $U\Lambda = \tilde{A}U$, we have $U_i \Lambda = \tilde{A}_i U = (\tilde{A}_i - \mathbb{E}A_i)U + (\mathbb{E}A_i)U$. We have

$$\begin{aligned} & \langle U_i \Lambda - \hat{\theta}_{z_i^*}, \hat{\theta}_a - \hat{\theta}_{z_i^*} \rangle \\ & = \langle (\tilde{A}_i - \mathbb{E}A_i)U, \hat{\theta}_a - \hat{\theta}_{z_i^*} \rangle + \langle (\mathbb{E}A_i)U - \hat{\theta}_{z_i^*}, \hat{\theta}_a - \hat{\theta}_{z_i^*} \rangle \\ & = (\tilde{A}_i - \mathbb{E}A_i)U \left(\hat{\theta}_a - \hat{\theta}_{z_i^*} \right)^T + \langle (\mathbb{E}A_i)U - \hat{\theta}_{z_i^*}, \hat{\theta}_a - \hat{\theta}_{z_i^*} \rangle \\ & = (A_i - \mathbb{E}A_i)U \left(\hat{\theta}_a - \hat{\theta}_{z_i^*} \right)^T + \langle (\mathbb{E}A_i)U - \hat{\theta}_{z_i^*}, \hat{\theta}_a - \hat{\theta}_{z_i^*} \rangle \\ & = (A_i - \mathbb{E}A_i)U^* \left(\theta_a^* - \theta_{z_i^*}^* \right)^T \\ & \quad + (A_i - \mathbb{E}A_i) \left(U \hat{\theta}_a^T - U^* \theta_a^{*T} - U \hat{\theta}_{z_i^*}^T + U^* \theta_{z_i^*}^{*T} \right) \\ & \quad + \langle (\mathbb{E}A_i)U - \hat{\theta}_{z_i^*}, \hat{\theta}_a - \hat{\theta}_{z_i^*} \rangle, \quad (34) \end{aligned}$$

where the third equation is by (17). Using Lemma 4, the first term in (34) can be simplified

$$\begin{aligned} & (A_i - \mathbb{E}A_i)U^* \left(\theta_a^* - \theta_{z_i^*}^* \right)^T \\ & = (p-q) \left(\sum_{j: z_j^* = a} (A_{ij} - \mathbb{E}A_{ij}) - \sum_{j: z_j^* = z_i^*, j \neq i} (A_{ij} - \mathbb{E}A_{ij}) \right). \end{aligned}$$

Together with (33) and (34), we have

$$\begin{aligned} & \mathbb{I}\left\{ \|U_i \Lambda - \hat{\theta}_a\| \leq \|U_i \Lambda - \hat{\theta}_{z_i^*}\| \right\} \mathbb{I}\{\mathcal{F}\} \\ & = \mathbb{I}\left\{ (p-q) \left(\sum_{j: z_j^* = a} (A_{ij} - \mathbb{E}A_{ij}) - \sum_{j: z_j^* = z_i^*, j \neq i} (A_{ij} - \mathbb{E}A_{ij}) \right) \right. \\ & \quad \left. + (A_i - \mathbb{E}A_i) \left(U \hat{\theta}_a^T - U^* \theta_a^{*T} - U \hat{\theta}_{z_i^*}^T + U^* \theta_{z_i^*}^{*T} \right) \right\} \mathbb{I}\{\mathcal{F}\} \end{aligned}$$

$$+ \left\langle (\mathbb{E}A_{i\cdot})U - \hat{\theta}_{z_i^*}, \hat{\theta}_a - \hat{\theta}_{z_i^*} \right\rangle \geq \frac{1}{2} \left\| \hat{\theta}_a - \hat{\theta}_{z_i^*} \right\|^2 \Big\} \mathbb{I}\{\mathcal{F}\}. \quad (35)$$

Next, we are going to give an upper bound for (35). Since $\mathbb{E}A_{i\cdot}$ and P_i only differ in their i th coordinates with $A_{ii} = 0$ and $P_{ii} = p$, we have

$$\begin{aligned} \left\| (\mathbb{E}A_{i\cdot})U - \hat{\theta}_{z_i^*} \right\| &= \left\| P_i U - P_{ii} U_{i\cdot} - \hat{\theta}_{z_i^*} \right\| \\ &\leq \left\| P_i U - \hat{\theta}_{z_i^*} \right\| + p \|U_{i\cdot}\| \\ &= \left\| \theta_{z_i^*}^* U^{*T} U - \hat{\theta}_{z_i^*} \right\| + p \|U_{i\cdot}\| \\ &= \left\| \theta_{z_i^*}^* U^{*T} U - \hat{\theta}_{z_i^*} U^T U \right\| + p \|U_{i\cdot}\| \\ &= \left\| \theta_{z_i^*}^* U^{*T} - \hat{\theta}_{z_i^*} U^T \right\| + p \|U_{i\cdot}\| \\ &\leq \max_{b \in [k]} \left\| U \hat{\theta}_b^T - U^* \theta_b^{*T} \right\| + p \\ &\leq C_2 \beta^{-0.5} k \sqrt{p} + p \\ &\leq 2C_2 \beta^{-0.5} k \sqrt{p}, \end{aligned}$$

where the second to last inequality is due to (14). We also have

$$\begin{aligned} &\left\| (A_{i\cdot} - \mathbb{E}A_{i\cdot}) \left(U \hat{\theta}_a^T - U^* \theta_a^{*T} - U \hat{\theta}_{z_i^*}^T + U^* \theta_{z_i^*}^{*T} \right) \right\| \\ &\leq \left\| (A_{i\cdot} - \mathbb{E}A_{i\cdot}) U^* U^{*T} \left(U \hat{\theta}_a^T - U^* \theta_a^{*T} - U \hat{\theta}_{z_i^*}^T + U^* \theta_{z_i^*}^{*T} \right) \right\| \\ &\quad + \left\| (A_{i\cdot} - \mathbb{E}A_{i\cdot}) (I_n - U^* U^{*T}) \right. \\ &\quad \times \left. \left(U \hat{\theta}_a^T - U^* \theta_a^{*T} - U \hat{\theta}_{z_i^*}^T + U^* \theta_{z_i^*}^{*T} \right) \right\| \\ &\leq 2 \left\| (A_{i\cdot} - \mathbb{E}A_{i\cdot}) U^* \right\| \max_{b \in [k]} \left\| U \hat{\theta}_b^T - U^* \theta_b^{*T} \right\| \\ &\quad + \left\| (A_{i\cdot} - \mathbb{E}A_{i\cdot}) (I_n - U^* U^{*T}) \left(U \hat{\theta}_a^T - U \hat{\theta}_{z_i^*}^T \right) \right\| \\ &\leq 2C_2 \beta^{-0.5} k \sqrt{p} \left\| (A_{i\cdot} - \mathbb{E}A_{i\cdot}) U^* \right\| \\ &\quad + \left\| (A_{i\cdot} - \mathbb{E}A_{i\cdot}) (I_n - U^* U^{*T}) U \right\| \left\| \hat{\theta}_a - \hat{\theta}_{z_i^*} \right\|, \end{aligned}$$

where in the last inequality we use (14) again. Then (35) leads to

$$\begin{aligned} &\mathbb{I}\left\{ \left\| U_{i\cdot} \Lambda - \hat{\theta}_a \right\| \leq \left\| U_{i\cdot} \Lambda - \hat{\theta}_{z_i^*} \right\| \right\} \mathbb{I}\{\mathcal{F}\} \\ &\leq \mathbb{I}\left\{ (p-q) \left(\sum_{j: z_j^* = a} (A_{ij} - \mathbb{E}A_{ij}) - \sum_{j: z_j^* = z_i^*, j \neq i} (A_{ij} - \mathbb{E}A_{ij}) \right) \right. \\ &\quad + 2C_2 \beta^{-0.5} k \sqrt{p} \left\| (A_{i\cdot} - \mathbb{E}A_{i\cdot}) U^* \right\| \\ &\quad + \left\| (A_{i\cdot} - \mathbb{E}A_{i\cdot}) (I_n - U^* U^{*T}) U \right\| \left\| \hat{\theta}_a - \hat{\theta}_{z_i^*} \right\| \\ &\quad \left. + 2C_2 \beta^{-0.5} k \sqrt{p} \left\| \hat{\theta}_a - \hat{\theta}_{z_i^*} \right\| \right\} \mathbb{I}\{\mathcal{F}\}. \quad (36) \end{aligned}$$

To further simplify the above display, we need to study $\left\| \hat{\theta}_a - \hat{\theta}_{z_i^*} \right\|$. From Lemma 1, we have $\left\| \theta_a^* - \theta_{z_i^*}^* \right\| = \sqrt{n_a + n_{z_i^*}}(p-q) \geq \sqrt{2\beta n/k}(p-q)$. Under the event \mathcal{F} , from (14), we have

$$\left\| \hat{\theta}_a - \hat{\theta}_{z_i^*} \right\|$$

$$\begin{aligned} &= \left\| \hat{\theta}_a U^T - \hat{\theta}_{z_i^*} U^T \right\| \\ &= \left\| \hat{\theta}_a U^T - \theta_a^* U^{*T} + \theta_a^* U^{*T} - \theta_{z_i^*}^* U^{*T} + \theta_{z_i^*}^* U^{*T} - \hat{\theta}_{z_i^*} U^T \right\| \\ &\geq \left\| \theta_a^* U^{*T} - \theta_{z_i^*}^* U^{*T} \right\| - 2 \max_{b \in [k]} \left\| \theta_b^* U^{*T} - \hat{\theta}_b U^T \right\| \\ &= \left\| \theta_a^* - \theta_{z_i^*}^* \right\| - 2 \max_{b \in [k]} \left\| \theta_b^* U^{*T} - \hat{\theta}_b U^T \right\| \\ &\geq \sqrt{n_a + n_{z_i^*}}(p-q) - 2C_2 \beta^{-0.5} k \sqrt{p}, \quad (37) \end{aligned}$$

which can be further lower bounded by $\frac{1}{2} \sqrt{\frac{2\beta n}{k}}(p-q)$ as $\frac{n(p-q)^2}{\beta^{-2} k^3 p} \rightarrow \infty$. Then,

$$\begin{aligned} &\frac{1}{2} \left\| \hat{\theta}_a - \hat{\theta}_{z_i^*} \right\|^2 - 2C_2 \beta^{-0.5} k \sqrt{p} \left\| \hat{\theta}_a - \hat{\theta}_{z_i^*} \right\| \\ &= \frac{1}{2} \left(1 - \frac{2C_2 \beta^{-0.5} k \sqrt{p}}{\left\| \hat{\theta}_a - \hat{\theta}_{z_i^*} \right\|} \right) \left\| \hat{\theta}_a - \hat{\theta}_{z_i^*} \right\|^2 \\ &\geq \frac{1}{2} \left(1 - \frac{2C_2 \beta^{-0.5} k \sqrt{p}}{\frac{1}{2} \sqrt{\frac{2\beta n}{k}}(p-q)} \right) \\ &\quad \times \left(\sqrt{n_a + n_{z_i^*}}(p-q) - 2C_2 \beta^{-0.5} k \sqrt{p} \right)^2 \\ &\geq \frac{1}{2} \left(1 - C_3 \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}} \right) (n_a + n_{z_i^*})(p-q)^2, \end{aligned}$$

for some constant $C_3 > 0$, where the last inequality holds as $\frac{n(p-q)^2}{\beta^{-2} k^3 p} \rightarrow \infty$. Similar to (37), we also have

$$\begin{aligned} \left\| \hat{\theta}_a - \hat{\theta}_{z_i^*} \right\| &\leq \left\| \theta_a^* - \theta_{z_i^*}^* \right\| + 2 \max_{b \in [k]} \left\| \theta_b^* U^{*T} - \hat{\theta}_b U^T \right\| \\ &\leq \sqrt{n_a + n_{z_i^*}}(p-q) + 2C_2 \beta^{-0.5} k \sqrt{p} \\ &\leq 1.1 \sqrt{n_a + n_{z_i^*}}(p-q), \end{aligned}$$

where the last inequality is under the assumption $\frac{n(p-q)^2}{\beta^{-2} k^3 p} \rightarrow \infty$. Then (36) becomes

$$\begin{aligned} &\mathbb{I}\left\{ \left\| U_{i\cdot} \Lambda - \hat{\theta}_a \right\| \leq \left\| U_{i\cdot} \Lambda - \hat{\theta}_{z_i^*} \right\| \right\} \mathbb{I}\{\mathcal{F}\} \\ &\leq \mathbb{I}\left\{ (p-q) \left(\sum_{j: z_j^* = a} (A_{ij} - \mathbb{E}A_{ij}) - \sum_{j: z_j^* = z_i^*, j \neq i} (A_{ij} - \mathbb{E}A_{ij}) \right) \right. \\ &\quad + 2C_2 \beta^{-0.5} k \sqrt{p} \left\| (A_{i\cdot} - \mathbb{E}A_{i\cdot}) U^* \right\| \\ &\quad + 1.1 \sqrt{n_a + n_{z_i^*}}(p-q) \left\| (A_{i\cdot} - \mathbb{E}A_{i\cdot}) (I_n - U^* U^{*T}) U \right\| \\ &\quad \left. \geq \frac{1}{2} \left(1 - C_3 \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}} \right) (n_a + n_{z_i^*})(p-q)^2 \right\} \mathbb{I}\{\mathcal{F}\}. \quad (38) \end{aligned}$$

We are going to split the indication function in (38) into several ones using the following fact: $\mathbb{I}\{x_1 + x_2 \geq x_3 + x_4\} \leq \mathbb{I}\{x_1 \geq x_3\} + \mathbb{I}\{x_2 \geq x_4\}$ for all $x_1, x_2, x_3, x_4 \in \mathbb{R}$. Consider a positive sequence $\rho = o(1)$ whose value will be determined later. We have

$$\begin{aligned} &\mathbb{I}\left\{ \left\| U_{i\cdot} \Lambda - \hat{\theta}_a \right\| \leq \left\| U_{i\cdot} \Lambda - \hat{\theta}_{z_i^*} \right\| \right\} \mathbb{I}\{\mathcal{F}\} \\ &\leq \mathbb{I}\left\{ \sum_{j: z_j^* = a} (A_{ij} - \mathbb{E}A_{ij}) - \sum_{j: z_j^* = z_i^*, j \neq i} (A_{ij} - \mathbb{E}A_{ij}) \geq \frac{1}{2} \times \right. \end{aligned}$$

$$\begin{aligned}
& \left(1 - 4\rho - C_3 \sqrt{\frac{\beta^{-2}k^3p}{n(p-q)^2}}\right) (n_a + n_{z_i^*})(p-q) \Big\} \mathbb{I}\{\mathcal{F}\} \\
& + \mathbb{I}\left\{\|(A_{i\cdot} - \mathbb{E}A_{i\cdot})U^*\| \geq \frac{\rho(n_a + n_{z_i^*})(p-q)^2}{2C_2\beta^{-0.5}k\sqrt{p}}\right\} \mathbb{I}\{\mathcal{F}\} \\
& + \mathbb{I}\{1.1 \|(A_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})U\| \\
& \quad \geq \rho\sqrt{n_a + n_{z_i^*}}(p-q)\} \mathbb{I}\{\mathcal{F}\} \\
& \leq \mathbb{I}\left\{\sum_{j:z_j^*=a} (A_{ij} - \mathbb{E}A_{ij}) - \sum_{j:z_j^*=z_i^*, j \neq i} (A_{ij} - \mathbb{E}A_{ij})\right. \\
& \quad \left. \geq \frac{1}{2} \left(1 - 4\rho - C_3 \sqrt{\frac{\beta^{-2}k^3p}{n(p-q)^2}}\right) (n_a + n_{z_i^*})(p-q)\right\} \\
& + \mathbb{I}\left\{\|(A_{i\cdot} - \mathbb{E}A_{i\cdot})U^*\| \geq \frac{\rho n(p-q)^2}{C_2\beta^{-1.5}k^2\sqrt{p}}\right\} \\
& + \mathbb{I}\left\{\|(A_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})U\| \right. \\
& \quad \left. \geq \rho\sqrt{\frac{\beta n}{k}}(p-q)\right\} \mathbb{I}\{\mathcal{F}\} \\
& =: G_{1,i,a} + G_{2,i} + G_{3,i} \tag{39}
\end{aligned}$$

where $G_{2,i}$ and $G_{3,i}$ do not depend on a .

To further decompose $G_{3,i}$, first we replace $A_{i\cdot}$ by $\tilde{A}_{i\cdot}$ as

$$\begin{aligned}
& (A_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})U \\
& = (\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})U \\
& \quad + (A_{i\cdot} - \tilde{A}_{i\cdot})(I_n - U^*U^{*T})U \\
& = (\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})U - (A_{i\cdot} - \tilde{A}_{i\cdot})U^*U^{*T}U,
\end{aligned}$$

where we use (17) in the last equation. Since $A_{ij} = \tilde{A}_{ij}$ for all $j \notin S$ and $\tilde{A}_{ij} = 0$ for all $j \in S$, we have $(A_{i\cdot} - \tilde{A}_{i\cdot})U^*U^{*T}U = \sum_{j \in S} A_{ij}U_j^*U^{*T}U$. As a result, $\|(A_{i\cdot} - \tilde{A}_{i\cdot})U^*U^{*T}U\| = \|\sum_{j \in S} A_{ij}U_j^*U^{*T}U\| \leq \|\sum_{j \in S} A_{ij}U_j^*\| \leq (\sum_{j \in S} A_{ij})\|U^*\|_{2,\infty} \leq \sqrt{\frac{k}{\beta n}} \sum_{j \in S} A_{ij}$ where the last inequality is due to (3). Hence,

$$\begin{aligned}
& \|(A_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})U\| \\
& \leq \|(\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})U\| + \sqrt{\frac{k}{\beta n}} \sum_{j \in S} A_{ij}.
\end{aligned}$$

Recall that \bar{U} and $\bar{U}^{(i)}$ are defined in (18) and (21) by applying $f_t(\cdot)$ to rows of U and $U^{(i)}$ where $t > 0$. For now, let us consider any $t > 0$. We will set $t = t_0$ later in the proof. Using $U = (U - \bar{U}) + (\bar{U} - \bar{U}^{(i)}W^{(i)}) + \bar{U}^{(i)}W^{(i)}$, we have

$$\begin{aligned}
& \|(A_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})U\| \\
& \leq \|(\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})(U - \bar{U})\| \\
& \quad + \|(\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})(\bar{U} - \bar{U}^{(i)}W^{(i)})\| \\
& \quad + \|(\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})\bar{U}^{(i)}W^{(i)}\| \\
& \quad + \sqrt{\frac{k}{\beta n}} \sum_{j \in S} A_{ij}.
\end{aligned}$$

The third term in the above display can be further simplified:

$$\begin{aligned}
& \|(\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})\bar{U}^{(i)}W^{(i)}\| \\
& = \|(\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})\bar{U}^{(i)}\| \\
& \leq \|(\tilde{A}_{i\cdot} - \tilde{P}_{i\cdot})(I_n - U^*U^{*T})\bar{U}^{(i)}\| \\
& \quad + \|(\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot}) - (\tilde{A}_{i\cdot} - \tilde{P}_{i\cdot})\| (I_n - U^*U^{*T})\bar{U}^{(i)}\| \\
& \leq \|(\tilde{A}_{i\cdot} - \tilde{P}_{i\cdot})(I_n - U^*U^{*T})\bar{U}^{(i)}\| \\
& \quad + \|(\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot}) - (\tilde{A}_{i\cdot} - \tilde{P}_{i\cdot})\|_1 \|(I_n - U^*U^{*T})\bar{U}^{(i)}\|_{2,\infty},
\end{aligned}$$

where the last inequality is by Hölder's inequality. Note that $(\tilde{A}_{ij} - \mathbb{E}A_{ij}) - (\tilde{A}_{ij} - \tilde{P}_{ij}) = (A_{ij} - \mathbb{E}A_{ij})\mathbb{I}\{j \in S^c \cap S_i\}$ for each $j \in [n]$. Together with (29), we have

$$\begin{aligned}
& \|(\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})\bar{U}^{(i)}W^{(i)}\| \\
& \leq \|(\tilde{A}_{i\cdot} - \tilde{P}_{i\cdot})(I_n - U^*U^{*T})\bar{U}^{(i)}\| \\
& \quad + \sum_{j \in S^c \cap S_i} |A_{ij} - \mathbb{E}A_{ij}| \left(t + \sqrt{\frac{k}{\beta n}}\right) \\
& \leq \|(\tilde{A}_{i\cdot} - \tilde{P}_{i\cdot})(I_n - U^*U^{*T})\bar{U}^{(i)}\| \\
& \quad + \sum_{j \in S_i} |A_{ij} - \mathbb{E}A_{ij}| \left(t + \sqrt{\frac{k}{\beta n}}\right).
\end{aligned}$$

Hence, the above displays lead to

$$\begin{aligned}
& \|(A_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})U\| \\
& \leq \|(\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})(U - \bar{U})\| \\
& \quad + \|(\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})(\bar{U} - \bar{U}^{(i)}W^{(i)})\| \\
& \quad + \|(\tilde{A}_{i\cdot} - \tilde{P}_{i\cdot})(I_n - U^*U^{*T})\bar{U}^{(i)}\| + \sqrt{\frac{k}{\beta n}} \sum_{j \in S} A_{ij} \\
& \quad + \left(t + \sqrt{\frac{k}{\beta n}}\right) \sum_{j \in S_i} |A_{ij} - \mathbb{E}A_{ij}|. \tag{40}
\end{aligned}$$

Then

$$\begin{aligned}
G_{3,i} & \leq \mathbb{I}\left\{\|(\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})(U - \bar{U})\| \right. \\
& \quad \left. \geq \frac{\rho}{4} \sqrt{\frac{\beta n}{k}}(p-q)\right\} \mathbb{I}\{\mathcal{F}\} \\
& + \mathbb{I}\left\{\|(\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^*U^{*T})(\bar{U} - \bar{U}^{(i)}W^{(i)})\| \right. \\
& \quad \left. \geq \frac{\rho}{4} \sqrt{\frac{\beta n}{k}}(p-q)\right\} \mathbb{I}\{\mathcal{F}\} \\
& + \mathbb{I}\left\{\|(\tilde{A}_{i\cdot} - \tilde{P}_{i\cdot})(I_n - U^*U^{*T})\bar{U}^{(i)}\| \right. \\
& \quad \left. \geq \frac{\rho}{4} \sqrt{\frac{\beta n}{k}}(p-q)\right\} \mathbb{I}\{\mathcal{F}\}
\end{aligned}$$

$$\begin{aligned}
& + \mathbb{I} \left\{ \sqrt{\frac{k}{\beta n}} \sum_{j \in S} A_{ij} \geq \frac{\rho}{8} \sqrt{\frac{\beta n}{k}} (p-q) \right\} \mathbb{I} \{\mathcal{F}\} \\
& + \mathbb{I} \left\{ \left(t + \sqrt{\frac{k}{\beta n}} \right) \sum_{j \in S_i} |A_{ij} - \mathbb{E}A_{ij}| \right. \\
& \quad \left. \geq \frac{\rho}{8} \sqrt{\frac{\beta n}{k}} (p-q) \right\} \mathbb{I} \{\mathcal{F}\} \\
& =: H_{1,i} + H_{2,i} + H_{3,i} + H_{4,i} + H_{5,i}.
\end{aligned} \tag{41}$$

So far, by (32), (39), and (41), we have

$$\begin{aligned}
& \frac{1}{n} \mathbb{E} \sum_{i \notin S} \mathbb{I} \{\hat{z}_i \neq z_i^*\} \mathbb{I} \{\mathcal{F}\} \\
& \leq \frac{1}{n} \mathbb{E} \sum_{i \notin S} \sum_{a \neq z_i^*} (G_{1,i,a} + G_{2,i} + G_{3,i}) \\
& \leq \frac{1}{n} \sum_{i \in [n]} \sum_{a \neq z_i^*} \mathbb{E} G_{1,i,a} + \frac{k}{n} \sum_{i \in [n]} \mathbb{E} G_{2,i} + \frac{k}{n} \sum_{i \in [n]} \mathbb{E} H_{1,i} \\
& \quad + \frac{k}{n} \sum_{i \in [n]} \mathbb{E} H_{2,i} + \frac{k}{n} \sum_{i \in [n]} (\mathbb{E} H_{3,i} + \mathbb{E} H_{4,i} + \mathbb{E} H_{5,i}).
\end{aligned} \tag{42}$$

We are going to analyze terms in (42) one by one. Consider any $i \in [n]$ and any $a \neq z_i^*$.

For $G_{1,i,a}$, we have

$$\begin{aligned}
& \mathbb{E} G_{1,i,a} \\
& \leq \mathbb{P} \left(\sum_{j: z_j^* = a} (A_{ij} - \mathbb{E}A_{ij}) - \sum_{j: z_j^* = z_i^*, j \neq i} (A_{ij} - \mathbb{E}A_{ij}) \right. \\
& \quad \left. \geq \frac{1}{2} \left(1 - 4\rho - C_3 \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}} \right) (n_a + n_{z_i^*} - 1)(p-q) \right) \\
& \leq \exp \left(-J_{n_a, n_{z_i^*} - 1, p, q} + \left(2\rho + \frac{C_3}{2} \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}} \right) \right. \\
& \quad \left. \times (n_a + n_{z_i^*} - 1) \frac{(p-q)^2}{q} \right) \\
& \leq \exp \left(-J_{n_a, n_{z_i^*}, p, q} + \frac{(p-q)^2}{4q} + \left(2\rho + \frac{C_3}{2} \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}} \right) \right. \\
& \quad \left. \times (n_a + n_{z_i^*} - 1) \frac{(p-q)^2}{q} \right) \\
& \leq \exp \left(- \left(1 - \frac{1}{n_{z_i^*} \frac{(p-q)^2}{8p}} \right) \left(\left(2\rho + \frac{C_3}{2} \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}} \right) \right. \right. \\
& \quad \left. \left. \times (n_a + n_{z_i^*} - 1) \frac{(p-q)^2}{q} + \frac{(p-q)^2}{4q} \right) \right) J_{n_a, n_{z_i^*}, p, q} \\
& \leq \exp \left(- \left(1 - \left(2\rho + \frac{C_3}{2} \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}} + \frac{1}{n} \right) \frac{16kp}{\beta q} \right) \right. \\
& \quad \left. \times J_{n_a, n_{z_i^*}, p, q} \right)
\end{aligned}$$

where the first inequality is due to Lemma 15 and the second and third inequalities are due to Lemma 13.

For $G_{2,i}$, we have $\|(A_{i\cdot} - \mathbb{E}A_{i\cdot})U^*\| \leq \sum_{j \in [k]} |(A_{i\cdot} - \mathbb{E}A_{i\cdot})U_j^*|$. Note that U_j^* is a unit vector with $\|U_j^*\|_\infty \leq \sqrt{k/(\beta n)}$ for each $j \in [k]$ according to Lemma 3. By the union bound and Bernstein inequality, we have

$$\begin{aligned}
& \mathbb{E} G_{2,i} \\
& \leq \sum_{j \in [k]} \mathbb{P} \left(|(A_{i\cdot} - \mathbb{E}A_{i\cdot})U_j^*| \geq \frac{\rho n(p-q)^2}{C_2 \beta^{-1.5} k^3 \sqrt{p}} \right) \\
& \leq 2 \sum_{j \in [k]} \exp \left(- \frac{\frac{1}{2} \left(\frac{\rho n(p-q)^2}{C_2 \beta^{-1.5} k^3 \sqrt{p}} \right)^2}{p \|U_j^*\|^2 + \frac{1}{3} \|U_j^*\|_\infty^2 \frac{\rho n(p-q)^2}{C_2 \beta^{-1.5} k^3 \sqrt{p}}} \right) \\
& \leq 2k \exp \left(- \frac{\frac{1}{2} \left(\frac{\rho n(p-q)^2}{C_2 \beta^{-1.5} k^3 \sqrt{p}} \right)^2}{p + \frac{1}{3} \sqrt{\frac{k}{\beta n}} \frac{\rho n(p-q)^2}{C_2 \beta^{-1.5} k^3 \sqrt{p}}} \right) \\
& \leq 2k \exp \left(- \left(\frac{1}{4} \frac{\rho^2 n^2 (p-q)^4}{C_2^2 \beta^{-3} k^6 p^2} \right) \wedge \left(\frac{3}{4} \frac{\sqrt{n p} \rho n(p-q)^2}{C_2 \beta^{-2} k^{3.5} p} \right) \right) \\
& \leq k \exp \left(- \frac{10n(p-q)^2}{p} \right),
\end{aligned}$$

where the last inequality holds as long as $\frac{\rho^2 n(p-q)^2}{\beta^{-3} k^6 p} \rightarrow \infty$ and $\frac{\rho \sqrt{n p}}{\beta^{-2} k^{3.5}} \rightarrow \infty$.

In the remaining part of the proof, we set $t = t_0$ whose value is given (13). We are going to analyze $H_{1,i}, H_{2,i}, H_{3,i}, H_{4,i}$, and $H_{5,i}$ term by term. For $H_{1,i}$, we have

$$\begin{aligned}
& \sum_{i \in [n]} H_{1,i} \\
& \leq \sum_{i \in [n]} \frac{16k}{\rho^2 \beta n(p-q)^2} \\
& \quad \times \left\| (\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^* U^{*T})(U - \bar{U}) \right\|^2 \mathbb{I} \{\mathcal{F}\} \\
& = \frac{16k}{\rho^2 \beta n(p-q)^2} \\
& \quad \times \left(\sum_{i \in [n]} \left\| (\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^* U^{*T})(U - \bar{U}) \right\|^2 \right) \mathbb{I} \{\mathcal{F}\} \\
& = \frac{16k}{\rho^2 \beta n(p-q)^2} \\
& \quad \times \left\| (\tilde{A} - \mathbb{E}A)(I_n - U^* U^{*T})(U - \bar{U}) \right\|_F^2 \mathbb{I} \{\mathcal{F}\} \\
& \leq \frac{16k}{\rho^2 \beta n(p-q)^2} \\
& \quad \times \left\| \tilde{A} - \mathbb{E}A \right\|^2 \left\| (I_n - U^* U^{*T})(U - \bar{U}) \right\|_F^2 \mathbb{I} \{\mathcal{F}\} \\
& \leq \frac{16C_0^2 k p}{\rho^2 \beta (p-q)^2} \|U - \bar{U}\|_F^2 \mathbb{I} \{\mathcal{F}\},
\end{aligned} \tag{43}$$

where in the last equation, we use the fact that $\{(\tilde{A}_{i\cdot} - \mathbb{E}A_{i\cdot})(I_n - U^* U^{*T})(U - \bar{U})\}_{i \in [n]}$ are rows of $(\tilde{A} - \mathbb{E}A)(I_n - U^* U^{*T})(U - \bar{U})$, and in the last inequality, we use $\|\tilde{A} - \mathbb{E}A\| \leq C_0 \sqrt{n p}$ under the event \mathcal{F} . By (18) and

Theorem 3, we have

$$\begin{aligned} \mathbb{E} \sum_{i \in [n]} H_{1,i} &\leq \frac{16C_0^2 kp}{\rho^2 \beta(p-q)^2} \mathbb{E} \sum_{i \in [n]} \|U_i\|^2 \mathbb{I} \{ \|U_i\| \mathbb{I} \{ \mathcal{F} \} \geq t_0 \} \\ &\leq \frac{16C_0^2 kp}{\rho^2 \beta(p-q)^2} t_0^2 n \exp \left(-\frac{3n(p-q)^2}{2p} \right) \\ &\leq nk \exp \left(-\frac{3n(p-q)^2}{2p} \right), \end{aligned}$$

where the second and the third inequalities hold under the assumption $\beta^{-1}, k = O(1), p \leq 1/2$, and $\frac{\rho^2 n(p-q)^2}{p} \rightarrow \infty$.

For $H_{2,i}$, using (22), we have

$$\begin{aligned} &\left\| (\tilde{A}_i - \mathbb{E}A_i)(I_n - U^*U^{*T})(\bar{U} - \bar{U}^{(i)}W^{(i)}) \right\| \\ &\leq \left\| \tilde{A}_i - \mathbb{E}A_i \right\| \left\| \bar{U} - \bar{U}^{(i)}W^{(i)} \right\| \\ &\leq \left\| \tilde{A} - \mathbb{E}A \right\| 6k^{1.5} \|U_i\|. \end{aligned} \quad (44)$$

Then

$$\begin{aligned} H_{2,i} &\leq \mathbb{I} \left\{ \left\| \tilde{A} - \mathbb{E}A \right\| 6k^{1.5} \|U_i\| \geq \frac{\rho}{4} \sqrt{\frac{\beta n}{k}}(p-q) \right\} \mathbb{I} \{ \mathcal{F} \} \\ &\leq \mathbb{I} \left\{ 6C_0 k^{1.5} \sqrt{np} \|U_i\| \geq \frac{\rho}{4} \sqrt{\frac{\beta n}{k}}(p-q) \right\} \mathbb{I} \{ \mathcal{F} \} \\ &\leq \mathbb{I} \{ \|U_i\| \geq t_0 \} \mathbb{I} \{ \mathcal{F} \}, \end{aligned}$$

where the last inequality holds under the assumption that $\beta^{-1}, k = O(1), p \leq 1/2$, and $\frac{\rho^2 n(p-q)^2}{p} \rightarrow \infty$. Then by Corollary 1, we have

$$\begin{aligned} \mathbb{E} \sum_{i \in [n]} H_{2,i} &\leq \mathbb{E} \sum_{i \in [n]} \mathbb{I} \{ \|U_i\| \geq t_0 \} \mathbb{I} \{ \mathcal{F} \} \\ &\leq n \exp \left(-\frac{3n(p-q)^2}{2p} \right). \end{aligned}$$

For $H_{3,i}$, using (27), we have

$$\begin{aligned} H_{3,i} &\leq \mathbb{I} \left\{ \left\| (\tilde{A}_i - \tilde{P}_i)(I_n - U^*U^{*T})\bar{U}^{(i)} \right\| \right. \\ &\quad \left. \geq \frac{\rho}{4} \sqrt{\frac{\beta n}{k}}(p-q) \right\} \mathbb{I} \{ \mathcal{F}_i \} \end{aligned}$$

and then

$$\begin{aligned} \mathbb{E} H_{3,i} &\leq \mathbb{E} \left(\mathbb{E} \left(\mathbb{I} \left\{ \left\| (\tilde{A}_i - \tilde{P}_i)(I_n - U^*U^{*T})\bar{U}^{(i)} \right\| \right. \right. \right. \\ &\quad \left. \left. \left. \geq \frac{\rho}{4} \sqrt{\frac{\beta n}{k}}(p-q) \right\} \mathbb{I} \{ \mathcal{F}_i \} \middle| S_i \cup \{i\} \right) \right). \end{aligned}$$

Note that according to (26) and (28), we have independence between $\tilde{A}_i - \tilde{P}_i$ and $\bar{U}^{(i)}$, \mathcal{F}_i when conditioned on $S_i \cup \{i\}$. On the other hand, we have (30) holds under the event \mathcal{F}_i and the assumption $\frac{n(p-q)^2}{\beta^{-2}k^2p} \rightarrow \infty$. Together with (29), we have

$$\begin{aligned} \mathbb{E} H_{3,i} &\leq \mathbb{E} \left(\sup_{\Delta \in \mathbb{R}^{n \times k}: \|\Delta\|_F \leq \frac{4(C_0+3)k\sqrt{kn}p}{\beta n(p-q)}, \|\Delta\|_{2,\infty} \leq 2t_0} \mathbb{E} \left(\mathbb{I} \left\{ \right. \right. \right. \end{aligned}$$

$$\begin{aligned} &\left. \left\| (\tilde{A}_i - \tilde{P}_i)\Delta \right\| \geq \frac{\rho}{4} \sqrt{\frac{\beta n}{k}}(p-q) \right\} \mathbb{I} \{ \mathcal{F}_i \} \middle| S_i \cup \{i\} \right) \\ &\leq \mathbb{E} \left(\sup_{\Delta \in \mathbb{R}^{n \times k}: \|\Delta\|_F \leq \frac{4(C_0+3)k\sqrt{kn}p}{\beta n(p-q)}, \|\Delta\|_{2,\infty} \leq 2t_0} \mathbb{E} \left(\mathbb{I} \left\{ \left\| (\tilde{A}_i - \tilde{P}_i)\Delta \right\| \geq \frac{\rho}{4} \sqrt{\frac{\beta n}{k}}(p-q) \right\} \middle| S_i \cup \{i\} \right) \right) \\ &\leq \mathbb{E} \left(\sup_{\Delta \in \mathbb{R}^{n \times k}: \|\Delta\|_F \leq \frac{4(C_0+3)k\sqrt{kn}p}{\beta n(p-q)}, \|\Delta\|_{2,\infty} \leq 2t_0} \mathbb{E} \sum_{l \in [k]} \sum_{\alpha \in \{-1,1\}} \left(\mathbb{I} \left\{ \alpha(\tilde{A}_i - \tilde{P}_i)\Delta_l \geq \frac{\rho}{8k} \sqrt{\frac{\beta n}{k}}(p-q) \right\} \middle| S_i \cup \{i\} \right) \right) \\ &\leq \sum_{l \in [k]} \sum_{\alpha \in \{-1,1\}} \mathbb{E} \left(\sup_{\Delta \in \mathbb{R}^{n \times k}: \|\Delta\|_F \leq \frac{4(C_0+3)k\sqrt{kn}p}{\beta n(p-q)}, \|\Delta\|_{2,\infty} \leq 2t_0} \mathbb{E} \left(\mathbb{I} \left\{ \alpha(\tilde{A}_i - \tilde{P}_i)\Delta_l \geq \frac{\rho}{8k} \sqrt{\frac{\beta n}{k}}(p-q) \right\} \middle| S_i \cup \{i\} \right) \right). \end{aligned}$$

Define

$$\mathcal{W} := \{w \in \mathbb{R}^n : \|w\| \leq \frac{4(C_0+3)k\sqrt{kn}p}{\beta n(p-q)}, \|w\|_\infty \leq 2t_0\}. \quad (45)$$

Then,

$$\begin{aligned} \mathbb{E} H_{3,i} &\leq \sum_{l \in [k]} \sum_{\alpha \in \{-1,1\}} \mathbb{E} \left(\sup_{w \in \mathcal{W}} \mathbb{E} \left(\mathbb{I} \left\{ \alpha(\tilde{A}_i - \tilde{P}_i)w \geq \frac{\rho}{8k} \sqrt{\frac{\beta n}{k}}(p-q) \right\} \middle| S_i \cup \{i\} \right) \right) \\ &= k \sum_{\alpha \in \{-1,1\}} \mathbb{E} \left(\sup_{w \in \mathcal{W}} \mathbb{E} \left(\mathbb{I} \left\{ \alpha(\tilde{A}_i - \tilde{P}_i)w \geq \frac{\rho}{8k} \sqrt{\frac{\beta n}{k}}(p-q) \right\} \middle| S_i \cup \{i\} \right) \right) \\ &= k \sum_{\alpha \in \{-1,1\}} \mathbb{E} \left(\sup_{w \in \mathcal{W}} \mathbb{P} \left(\alpha \sum_{j \in S_i^c} (A_{ij} - \mathbb{E}A_{ij})w_j \geq \frac{\rho}{8k} \sqrt{\frac{\beta n}{k}}(p-q) \middle| S_i \cup \{i\} \right) \right), \end{aligned} \quad (46)$$

where the last equation is due to the definition of \tilde{A}_i and \tilde{P}_i in (25). Now consider any $\alpha \in \{-1,1\}$ and any $w \in \mathcal{W}$. Then by Chernoff bound, the conditional probability in the above display can be upper bounded by

$$\begin{aligned} &\log \mathbb{P} \left(\alpha \sum_{j \in S_i^c} (A_{ij} - \mathbb{E}A_{ij})w_j \geq \frac{\rho}{8k} \sqrt{\frac{\beta n}{k}}(p-q) \middle| S_i \cup \{i\} \right) \\ &\leq -s \frac{\rho}{8k} \sqrt{\frac{\beta n}{k}}(p-q) \end{aligned}$$

$$\begin{aligned}
& + \log \mathbb{E} \exp \left(s \alpha \sum_{j \in S_i^c} (A_{ij} - \mathbb{E} A_{ij}) w_j \middle| S_i \cup \{i\} \right) \\
& = -s \frac{\rho}{8k} \sqrt{\frac{\beta n}{k}} (p-q) + \log \prod_{j \in S_i^c} \mathbb{E} \exp (s \alpha (A_{ij} - \mathbb{E} A_{ij}) w_j) \\
& = -s \frac{\rho}{8k} \sqrt{\frac{\beta n}{k}} (p-q) + \sum_{j \in S_i^c} \log \mathbb{E} \exp (s \alpha (A_{ij} - \mathbb{E} A_{ij}) w_j) \\
& \leq -s \frac{\rho}{8k} \sqrt{\frac{\beta n}{k}} (p-q) + p s^2 \|w\|^2 \exp (s \|w\|_\infty) \\
& \leq -s \frac{\rho}{8k} \sqrt{\frac{\beta n}{k}} (p-q) \\
& \quad + p s^2 \left(\frac{4(C_0 + 3)k\sqrt{kn}p}{\beta n(p-q)} \right)^2 \exp (2st_0), \tag{47}
\end{aligned}$$

for any $s > 0$, where the second to last inequality is due to Lemma 10 under the assumption $p \leq 1/2$. Choose $s = \frac{8k}{\rho} \sqrt{\frac{k}{\beta n} \frac{4n(p-q)}{p}}$ so that the first term in the above display is equal to $-\frac{4n(p-q)^2}{p}$. Then, we have

$$\begin{aligned}
& \log \mathbb{P} \left(\alpha \sum_{j \in S_i^c} (A_{ij} - \mathbb{E} A_{ij}) w_j \geq \frac{\rho}{8k} \sqrt{\frac{\beta n}{k}} (p-q) \middle| S_i \cup \{i\} \right) \\
& \leq -\frac{4n(p-q)^2}{p} \\
& \quad + p \left(\frac{8k}{\rho} \sqrt{\frac{k}{\beta n} \frac{4n(p-q)}{p}} \frac{4(C_0 + 3)k\sqrt{kn}p}{\beta n(p-q)} \right)^2 \\
& \quad \times \exp \left(\frac{8k}{\rho} \sqrt{\frac{k}{\beta n} \frac{4n(p-q)}{p}} 2t_0 \right) \\
& = -\frac{4n(p-q)^2}{p} \\
& \quad + \frac{128^2(C_0 + 3)^2 k^6}{\rho^2 \beta^3} \exp \left(\frac{64 \times 160 k^4 (p-q)}{\rho \beta^2 p} \right),
\end{aligned}$$

where in the last equation we use the definition of t_0 in (13). When $\beta^{-1}, k = O(1)$ and $\frac{n(p-q)^2}{p} \rightarrow \infty$, a sufficient condition for the second term in the above display to be dominated by the first term is that ρ satisfies $\rho^{-1} = o\left(\log\left(\frac{n(p-q)^2}{p}\right)\right)$. Then

$$\begin{aligned}
& \log \mathbb{P} \left(\alpha \sum_{j \in S_i^c} (A_{ij} - \mathbb{E} A_{ij}) w_j \geq \frac{\rho}{8k} \sqrt{\frac{\beta n}{k}} (p-q) \middle| S_i \cup \{i\} \right) \\
& \leq -\frac{2n(p-q)^2}{p}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathbb{E} H_{3,i} & \leq k \sum_{\alpha \in \{-1,1\}} \mathbb{E} \exp \left(-\frac{2n(p-q)^2}{p} \right) \\
& \leq 2k \exp \left(-\frac{2n(p-q)^2}{p} \right).
\end{aligned}$$

For $H_{4,i}$, we first have $\sum_{j \in S} A_{ij} \leq \sum_{j \in S_i} A_{ij}$ due to (24) and the fact $A_{ii} = 0$. Note that conditioned on S_i , $\sum_{j \in S_i} A_{ij} \sim \text{Binomial}(|S_i|, p)$. This leads to

$\mathbb{E} \sum_{j \in S_i} A_{ij} = p \mathbb{E} |S_i| \leq np \exp(-128np)$ where the last inequality is by Lemma 6. Then

$$\begin{aligned}
\mathbb{E} H_{4,i} & \leq \mathbb{E} \mathbb{I} \left\{ \sum_{j \in S_i} A_{ij} \geq \frac{\rho \beta n (p-q)}{8k} \right\} \\
& \leq \frac{8k}{\rho \beta n (p-q)} \mathbb{E} \sum_{j \in S_i} A_{ij} \\
& \leq \frac{8kp}{\rho \beta (p-q)} \exp(-128np) \\
& = 8 \sqrt{\frac{p}{\rho^2 \beta^2 n (p-q)^2}} k \sqrt{np} \exp(-128np) \\
& \leq k \exp(-64np),
\end{aligned}$$

where the last inequality holds under the assumption that $\frac{\rho^2 n(p-q)^2}{\beta^{-2} p} \rightarrow \infty$. Similarly, for $H_{5,i}$, we have

$$\begin{aligned}
\mathbb{E} H_{5,i} & \leq \mathbb{E} \mathbb{I} \left\{ 2t_0 \sum_{j \in S_i} |A_{ij} - \mathbb{E} A_{ij}| \geq \frac{\rho}{8} \sqrt{\frac{\beta n}{k}} (p-q) \right\} \\
& \leq \frac{2t_0}{\frac{\rho}{8} \sqrt{\frac{\beta n}{k}} (p-q)} \mathbb{E} \sum_{j \in S_i} |A_{ij} - \mathbb{E} A_{ij}| \\
& \leq \frac{2t_0}{\frac{\rho}{8} \sqrt{\frac{\beta n}{k}} (p-q)} 2p \mathbb{E} |S_i| \\
& \leq \frac{2t_0}{\frac{\rho}{8} \sqrt{\frac{\beta n}{k}} (p-q)} 2np \exp(-128np) \\
& \leq k \exp(-64np),
\end{aligned}$$

where the last inequality holds under the assumption that $\beta^{-1}, k = O(1)$, and $\frac{\rho^2 n(p-q)^2}{p} \rightarrow \infty$.

Now we can combine the above results together to derive the final conclusion. From (42), we have

$$\begin{aligned}
& \frac{1}{n} \mathbb{E} \sum_{i \notin S} \mathbb{I} \{ \hat{z}_i \neq z_i^* \} \mathbb{I} \{ \mathcal{F} \} \\
& \leq \frac{1}{n} \sum_{i \in S} \sum_{a \neq z_i^*} \exp \left(- \left(1 - \left(2\rho + \frac{C_3}{2} \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}} + \frac{1}{n} \right) \right. \right. \\
& \quad \times \left. \left. \frac{16kp}{\beta q} \right) J_{n_a, n_{z_i^*}, p, q} \right) \\
& \quad + \frac{k}{n} \sum_{i \in [n]} k \exp \left(-\frac{10n(p-q)^2}{p} \right) \\
& \quad + \frac{k}{n} n \exp \left(-\frac{3n(p-q)^2}{2p} \right) \\
& \quad + \frac{k}{n} \sum_{i \in [n]} \left(2k \exp \left(-\frac{2n(p-q)^2}{p} \right) + 2k \exp(-64np) \right) \\
& \leq k \exp \left(- \left(1 - \left(2\rho + \frac{C_3}{2} \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}} + \frac{1}{n} \right) \frac{16kp}{\beta q} \right) \right. \\
& \quad \times \left. \min_{1 \leq a \neq b \leq k} J_{n_a, n_b, p, q} \right) + 4k^2 \exp \left(-\frac{3n(p-q)^2}{2p} \right).
\end{aligned}$$

So far, we require $\beta^{-1}, k = O(1)$, $0 < q < p \leq 1/2$, $\frac{n(p-q)^2}{p} \rightarrow \infty$, and $\rho^{-1} = o\left(\log\left(\frac{n(p-q)^2}{p}\right)\right)$. Then using (31), we have

$$\begin{aligned} \mathbb{E}\ell(\hat{z}, z^*) &\leq k \exp\left(-\left(1 - \left(2\rho + \frac{C_3}{2} \sqrt{\frac{\beta^{-2}k^3p}{n(p-q)^2}} + \frac{1}{n}\right) \frac{16kp}{\beta q}\right)\right. \\ &\quad \times \min_{1 \leq a \neq b \leq k} J_{n_a, n_b, p, q} \Bigg) \\ &\quad + 4k^2 \exp\left(-\frac{3n(p-q)^2}{2p}\right) + \exp(-128np) + 2n^{-3}. \end{aligned} \quad (48)$$

When $p \leq 1/10$ is further assumed, by Lemma 13, we have $J_{\min} = \min_{1 \leq a \neq b \leq k} J_{n_a, n_b, p, q} \leq \max_{a \in [k]} n_a 4(p-q)^2/(3p) \leq 4n(p-q)^2/(3p)$. As a result, the exponents $128np$ and $\frac{3n(p-q)^2}{2p}$ in (48) are bigger than J_{\min} . We can let $\rho^{-1} = \left(\log\left(\frac{n(p-q)^2}{p}\right)\right)^{\frac{1}{2}}$. With $\frac{p}{q} = O(1)$ further assumed, there exists some constant C_4 such that

$$\begin{aligned} \mathbb{E}\ell(\hat{z}, z^*) \mathbb{I}\{\mathcal{F}\} &\leq \\ &\exp\left(-\left(1 - C_4 \left(\log\left(\frac{n(p-q)^2}{p}\right)\right)^{-\frac{1}{2}}\right) J_{\min}\right) + 2n^{-3}. \end{aligned}$$

B. Proof of Theorem 3

We first have a decomposition of rows of U . Consider any $i \in [n]$. Due to $U\Lambda = \tilde{A}U$, we have $U = \tilde{A}U\Lambda^{-1}$ and its i th row satisfies $U_i = \tilde{A}_i U \Lambda^{-1}$. Then we have

$$U_i = \tilde{A}_i U \Lambda^{-1} = (\mathbb{E}A_i)U \Lambda^{-1} + (\tilde{A}_i - \mathbb{E}A_i)U \Lambda^{-1}.$$

This gives us

$$\begin{aligned} \|U_i\| \mathbb{I}\{\mathcal{F}\} &\leq \|(\mathbb{E}A_i)U \Lambda^{-1}\| \mathbb{I}\{\mathcal{F}\} \\ &\quad + \|(\tilde{A}_i - \mathbb{E}A_i)U\| \|\Lambda^{-1}\| \mathbb{I}\{\mathcal{F}\}. \end{aligned} \quad (49)$$

For the term $\|(\mathbb{E}A_i)U \Lambda^{-1}\|$ in (49), we have

$$\begin{aligned} (\mathbb{E}A_i)U \Lambda^{-1} &= P_i U \Lambda^{-1} - p U_i \Lambda^{-1} \\ &= U_i^* \Lambda^* U^{*T} U \Lambda^{-1} - p U_i \Lambda^{-1} \\ &= U_i^* U^{*T} U + U_i^* (\Lambda^* U^{*T} U - U^{*T} U \Lambda) \Lambda^{-1} \\ &\quad - p U_i \Lambda^{-1}. \end{aligned}$$

Hence,

$$\begin{aligned} \|(\mathbb{E}A_i)U \Lambda^{-1}\| &\leq \|U_i^*\| + \|U_i^*\| \|\Lambda^* U^{*T} U - U^{*T} U \Lambda\| \|\Lambda^{-1}\| \\ &\quad + p \|U_i\| \|\Lambda^{-1}\|. \end{aligned}$$

Note that $\|U_i\| \leq \|U\| = 1$. From Lemma 3, we have $\|U_i^*\| \leq 1/\sqrt{\beta n/k}$. Note that $\Lambda^* U^{*T} = U^{*T} P$ and $U \Lambda = \tilde{A} U^T$. We have $\Lambda^* U^{*T} U - U^{*T} U \Lambda = U^{*T} P U - U^{*T} \tilde{A} U^T = U^{*T} (P - \tilde{A}) U$ and thus $\|\Lambda^* U^{*T} U - U^{*T} U \Lambda\| \leq \|\tilde{A} - P\|$. By Lemma 5, we have $\|\Lambda^{-1}\| \leq 2k/(\beta n(p-q))$. As a result,

$$\|(\mathbb{E}A_i)U \Lambda^{-1}\| \mathbb{I}\{\mathcal{F}\}$$

$$\begin{aligned} &\leq \sqrt{\frac{k}{\beta n}} + \sqrt{\frac{k}{\beta n}} C_0 \sqrt{np} \frac{2k}{\beta n(p-q)} + p \frac{2k}{\beta n(p-q)} \\ &\leq 2\sqrt{\frac{k}{\beta n}}, \end{aligned}$$

where the last inequality holds under the assumption that $\frac{n(p-q)^2}{\beta^{-2}k^2p} \rightarrow \infty$. Then (49) leads to

$$\|U_i\| \mathbb{I}\{\mathcal{F}\} \leq 2\sqrt{\frac{k}{\beta n}} + \frac{4k}{\beta n(p-q)} \|(\tilde{A}_i - \mathbb{E}A_i)U\| \mathbb{I}\{\mathcal{F}\}. \quad (50)$$

From (16), we have

$$\sum_{i \in [n]} \|U_i\|^2 \mathbb{I}\{\|U_i\| \geq t\} = \sum_{i \notin S} \|U_i\|^2 \mathbb{I}\{\|U_i\| \geq t\}. \quad (51)$$

Hence, we only need to consider indexes not in S . From now on, consider any $i \notin S$. Then together with (17), (50) leads to

$$\|U_i\| \mathbb{I}\{\mathcal{F}\} \leq 2\sqrt{\frac{k}{\beta n}} + \frac{4k}{\beta n(p-q)} \|(A_i - \mathbb{E}A_i)U\| \mathbb{I}\{\mathcal{F}\}. \quad (52)$$

We are going to decompose $(A_i - \mathbb{E}A_i)U$ in a similar way as in the proof of Theorem 2. Recall that \bar{U} and $\bar{U}^{(i)}$ are defined in (18) and (21) by applying $f_t(\cdot)$ to rows of U and $U^{(i)}$ where $t > 0$. For any $t > 0$, we have

$$\begin{aligned} &\|(A_i - \mathbb{E}A_i)U\| \\ &\leq \|(A_i - \mathbb{E}A_i)U^* U^{*T} U\| \\ &\quad + \|(A_i - \mathbb{E}A_i)(I - U^* U^{*T})U\| \\ &\leq \|(A_i - \mathbb{E}A_i)U^*\| + \|(A_i - \mathbb{E}A_i)(I - U^* U^{*T})U\| \\ &\leq \|(A_i - \mathbb{E}A_i)U^*\| \\ &\quad + \|(\tilde{A}_i - \mathbb{E}A_i)(I_n - U^* U^{*T})(U - \bar{U})\| \\ &\quad + \|(\tilde{A}_i - \mathbb{E}A_i)(I_n - U^* U^{*T})(\bar{U} - \bar{U}^{(i)} W^{(i)})\| \\ &\quad + \|(\tilde{A}_i - \tilde{P}_i)(I_n - U^* U^{*T})\bar{U}^{(i)}\| + \sqrt{\frac{k}{\beta n}} \sum_{j \in S} A_{ij} \\ &\quad + \left(t + \sqrt{\frac{k}{\beta n}}\right) \sum_{j \in S_i} |A_{ij} - \mathbb{E}A_{ij}|, \end{aligned}$$

where the last inequality is by (40). From (44), we have $\|(\tilde{A}_i - \mathbb{E}A_i)(I_n - U^* U^{*T})(\bar{U} - \bar{U}^{(i)} W^{(i)})\| \leq 6C_0 k^{1.5} \sqrt{np} \|U_i\|$ under the event \mathcal{F} . Then (52) gives

$$\begin{aligned} &\|U_i\| \mathbb{I}\{\mathcal{F}\} \\ &\leq \left(2\sqrt{\frac{k}{\beta n}} + \frac{4k}{\beta n(p-q)}\right) \|(A_i - \mathbb{E}A_i)U^*\| \\ &\quad + \frac{4k}{\beta n(p-q)} \|(\tilde{A}_i - \mathbb{E}A_i)(I_n - U^* U^{*T})(U - \bar{U})\| \\ &\quad + \frac{24C_0 k^{2.5} \sqrt{np}}{\beta n(p-q)} \|U_i\| \\ &\quad + \frac{4k}{\beta n(p-q)} \|(\tilde{A}_i - \tilde{P}_i)(I_n - U^* U^{*T})\bar{U}^{(i)}\| \end{aligned}$$

$$\begin{aligned}
& + \frac{4k}{\beta n(p-q)} \sqrt{\frac{k}{\beta n}} \sum_{j \in S} A_{ij} \\
& + \frac{4k}{\beta n(p-q)} \left(t + \sqrt{\frac{k}{\beta n}} \right) \sum_{j \in S_i} |A_{ij} - \mathbb{E}A_{ij}| \mathbb{I}\{\mathcal{F}\}.
\end{aligned}$$

Under the assumption that $\frac{n(p-q^2)}{\beta^{-2}k^5p} \rightarrow \infty$, we have $\left(1 - \frac{24C_0k^{2.5}\sqrt{p}}{\beta\sqrt{n(p-q)}}\right) > \frac{1}{2}$. After arrangement, we have

$$\begin{aligned}
& \|U_i\| \mathbb{I}\{\mathcal{F}\} \\
& \leq \left(4\sqrt{\frac{k}{\beta n}} + \frac{8k}{\beta n(p-q)} \|(A_i - \mathbb{E}A_i)U^*\| \right. \\
& \quad + \frac{8k}{\beta n(p-q)} \|(\tilde{A}_i - \mathbb{E}A_i)(I_n - U^*U^{*T})(U - \bar{U})\| \\
& \quad + \frac{8k}{\beta n(p-q)} \|(\tilde{A}_i - \check{P}_i)(I_n - U^*U^{*T})\bar{U}^{(i)}\| \\
& \quad + \frac{8k}{\beta n(p-q)} \sqrt{\frac{k}{\beta n}} \sum_{j \in S} A_{ij} \\
& \quad \left. + \frac{8k}{\beta n(p-q)} \left(t + \sqrt{\frac{k}{\beta n}} \right) \sum_{j \in S_i} |A_{ij} - \mathbb{E}A_{ij}| \right) \mathbb{I}\{\mathcal{F}\}.
\end{aligned}$$

By Lemma 11, we have

$$\begin{aligned}
& \|U_i\|^2 \mathbb{I}\{\|U_i\| \geq t\} \mathbb{I}\{\mathcal{F}\} \\
& \leq 25 \left(\left(4\sqrt{\frac{k}{\beta n}} \right)^2 \mathbb{I}\left\{ 4\sqrt{\frac{k}{\beta n}} \geq \frac{t}{5} \right\} \right. \\
& \quad + \left(\frac{8k}{\beta n(p-q)} \|(A_i - \mathbb{E}A_i)U^*\| \right)^2 \\
& \quad \times \mathbb{I}\left\{ \frac{8k}{\beta n(p-q)} \|(A_i - \mathbb{E}A_i)U^*\| \geq \frac{t}{5} \right\} \\
& \quad + \left(\frac{8k \|(\tilde{A}_i - \mathbb{E}A_i)(I_n - U^*U^{*T})(U - \bar{U})\|}{\beta n(p-q)} \right)^2 \\
& \quad \times \mathbb{I}\left\{ \frac{8k \|(\tilde{A}_i - \mathbb{E}A_i)(I_n - U^*U^{*T})(U - \bar{U})\|}{\beta n(p-q)} \geq \frac{t}{5} \right\} \\
& \quad + \left(\frac{8k \|(\tilde{A}_i - \check{P}_i)(I_n - U^*U^{*T})\bar{U}^{(i)}\|}{\beta n(p-q)} \right)^2 \\
& \quad \times \mathbb{I}\left\{ \frac{8k \|(\tilde{A}_i - \check{P}_i)(I_n - U^*U^{*T})\bar{U}^{(i)}\|}{\beta n(p-q)} \geq \frac{t}{5} \right\} \\
& \quad + \left(\frac{8k}{\beta n(p-q)} \sqrt{\frac{k}{\beta n}} \sum_{j \in S} A_{ij} \right)^2 \\
& \quad \times \mathbb{I}\left\{ \frac{8k}{\beta n(p-q)} \sqrt{\frac{k}{\beta n}} \sum_{j \in S} A_{ij} \geq \frac{t}{5} \right\} \\
& \quad + \left(\frac{8k}{\beta n(p-q)} \left(t + \sqrt{\frac{k}{\beta n}} \right) \sum_{j \in S_i} |A_{ij} - \mathbb{E}A_{ij}| \right)^2 \mathbb{I}\left\{ \right.
\end{aligned}$$

$$\frac{8k}{\beta n(p-q)} \left(t + \sqrt{\frac{k}{\beta n}} \right) \sum_{j \in S_i} |A_{ij} - \mathbb{E}A_{ij}| \geq \frac{t}{5} \Bigg\} \mathbb{I}\{\mathcal{F}\}.$$

For any $t > 20\sqrt{\frac{k}{\beta n}}$, we have $\left(4\sqrt{\frac{k}{\beta n}}\right)^2 \mathbb{I}\left\{4\sqrt{\frac{k}{\beta n}} \geq \frac{t}{5}\right\} = 0$. With this and the fact that an indication function is always smaller or equal to 1, the above display can be simplified into

$$\begin{aligned}
& \|U_i\|^2 \mathbb{I}\{\|U_i\| \geq t\} \mathbb{I}\{\mathcal{F}\} \\
& \leq 25 \left(\frac{8k}{\beta n(p-q)} \|(A_i - \mathbb{E}A_i)U^*\| \right)^2 \\
& \quad \times \mathbb{I}\left\{ \frac{8k}{\beta n(p-q)} \|(A_i - \mathbb{E}A_i)U^*\| \geq \frac{t}{5} \right\} \\
& \quad + 25 \left(\frac{8k}{\beta n(p-q)} \right)^2 \\
& \quad \times \left\| (\tilde{A}_i - \mathbb{E}A_i)(I_n - U^*U^{*T})(U - \bar{U}) \right\|^2 \mathbb{I}\{\mathcal{F}\} \\
& \quad + 25 \left(\frac{8k}{\beta n(p-q)} \left\| (\tilde{A}_i - \check{P}_i)(I_n - U^*U^{*T})\bar{U}^{(i)} \right\| \right)^2 \\
& \quad \times \mathbb{I}\left\{ \frac{8k}{\beta n(p-q)} \left\| (\tilde{A}_i - \check{P}_i)(I_n - U^*U^{*T})\bar{U}^{(i)} \right\| \right. \\
& \quad \left. \geq \frac{t}{5} \right\} \mathbb{I}\{\mathcal{F}\} \\
& \quad + 25 \left(\frac{8k}{\beta n(p-q)} \right)^2 \frac{k}{\beta n} \left(\sum_{j \in S} A_{ij} \right)^2 \\
& \quad + 25 \left(\frac{16k}{\beta n(p-q)} \right)^2 t^2 \left(\sum_{j \in S_i} |A_{ij} - \mathbb{E}A_{ij}| \right)^2 \\
& =: H'_{0,i} + H'_{1,i} + H'_{3,i} + H'_{4,i} + H'_{5,i}.
\end{aligned}$$

Note the similarity between $H'_{1,i}$, $H'_{3,i}$, $H'_{4,i}$, $H'_{5,i}$ and $H_{1,i}$, $H_{3,i}$, $H_{4,i}$, $H_{5,i}$ defined in the proof of Theorem 2. Summing over all $i \notin S$, we have

$$\begin{aligned}
& \sum_{i \notin S} \|U_i\|^2 \mathbb{I}\{\|U_i\| \geq t\} \\
& \leq \sum_{i \in [n]} H'_{0,i} + \sum_{i \in [n]} H'_{1,i} + \sum_{i \in [n]} H'_{3,i} + \sum_{i \in [n]} H'_{4,i} + \sum_{i \in [n]} H'_{5,i}.
\end{aligned} \tag{53}$$

For $\sum_{i \in [n]} H'_{1,i}$, following the analysis to establish (43) for $\sum_{i \in [n]} H_{1,i}$ in the proof of Theorem 2, we have

$$\begin{aligned}
& \sum_{i \in [n]} H'_{1,i} \\
& \leq 25 \left(\frac{8k}{\beta n(p-q)} \right)^2 \\
& \quad \times \sum_{i \in [n]} \left\| (\tilde{A}_i - \mathbb{E}A_i)(I_n - U^*U^{*T})(U - \bar{U}) \right\|^2 \mathbb{I}\{\mathcal{F}\} \\
& \leq 25 \left(\frac{8k}{\beta n(p-q)} \right)^2 (C_0\sqrt{np})^2 \|U - \bar{U}\|_F^2 \mathbb{I}\{\mathcal{F}\} \\
& \leq \frac{40^2 C_0^2 k^2 p}{\beta^2 n(p-q)^2} \|U - \bar{U}\|_F^2 \mathbb{I}\{\mathcal{F}\}.
\end{aligned}$$

Using (18) and (16), we have

$$\begin{aligned} \sum_{i \in [n]} H'_{1,i} &\leq \frac{40^2 C_0^2 k^2 p}{\beta^2 n(p-q)^2} \sum_{i \in [n]} \|U_{i \cdot}\|^2 \mathbb{I}\{\|U_{i \cdot}\| \geq t\} \\ &= \frac{40^2 C_0^2 k^2 p}{\beta^2 n(p-q)^2} \sum_{i \notin S} \|U_{i \cdot}\|^2 \mathbb{I}\{\|U_{i \cdot}\| \geq t\}. \end{aligned}$$

Under the assumption $\frac{n(p-q)^2}{\beta^{-2} k^2 p} \rightarrow \infty$, we have $\left(1 - \frac{40^2 C_0^2 k^2 p}{\beta^2 n(p-q)^2}\right) \geq \frac{1}{2}$. Rearranging (53), we have

$$\begin{aligned} &\sum_{i \notin S} \|U_{i \cdot}\|^2 \mathbb{I}\{\|U_{i \cdot}\| \geq t\} \\ &\leq 2 \sum_{i \in [n]} H'_{0,i} + 2 \sum_{i \in [n]} H'_{3,i} + 2 \sum_{i \in [n]} H'_{4,i} + 2 \sum_{i \in [n]} H'_{5,i}. \end{aligned}$$

Using (51), we have

$$\begin{aligned} &\mathbb{E} \sum_{i \in [n]} \|U_{i \cdot}\|^2 \mathbb{I}\{\|U_{i \cdot}\| \geq t\} \\ &\leq 2\mathbb{E} \sum_{i \in [n]} H'_{0,i} + 2\mathbb{E} \sum_{i \in [n]} H'_{3,i} + 2\mathbb{E} \sum_{i \in [n]} H'_{4,i} + 2\mathbb{E} \sum_{i \in [n]} H'_{5,i}. \end{aligned} \quad (54)$$

In the following, we are going to consider any $i \in [n]$ and analyze $\mathbb{E}H'_{0,i}$, $\mathbb{E}H'_{3,i}$, $\mathbb{E}H'_{4,i}$, and $\mathbb{E}H'_{5,i}$, term by term. We first analyze the latter two as their analysis is more straightforward. Similar to the analysis for $H_{4,i}$ in the proof of Theorem 2, we have

$$\begin{aligned} \mathbb{E} \left(\sum_{j \in S} A_{ij} \right)^2 &\leq \mathbb{E} \left(\sum_{j \in S_i} A_{ij} \right)^2 \\ &= \mathbb{E} \left(\mathbb{E} \left(\left(\sum_{j \in S_i} A_{ij} \right)^2 \middle| S_i \right) \right) \\ &\leq \mathbb{E} ((|S_i|p)^2 + |S_i|p) \\ &\leq (np+1) \mathbb{E} |S_i|p \\ &\leq 2(np)^2 \exp(-128np), \end{aligned}$$

where the last inequality is by Lemma 6. Then,

$$\begin{aligned} \mathbb{E}H'_{4,i} &= 25 \left(\frac{8k}{\beta n(p-q)} \right)^2 \frac{k}{\beta n} \mathbb{E} \left(\sum_{j \in S} A_{ij} \right)^2 \\ &\leq 2 \frac{40^2 k^3 p^2}{\beta^3 n(p-q)^2} \exp(-128np). \end{aligned}$$

Similarly, we have

$$\begin{aligned} \mathbb{E}H'_{5,i} &\leq 25 \left(\frac{16k}{\beta n(p-q)} \right)^2 t^2 \mathbb{E} \left(\sum_{j \in S_i} |A_{ij} - \mathbb{E}A_{ij}| \right)^2 \\ &\leq 25 \left(\frac{16k}{\beta n(p-q)} \right)^2 t^2 \mathbb{E} ((2p|S_i|)^2 + 2p|S_i|) \\ &\leq 6 \frac{160k^2 t^2 np^2}{\beta^2 n(p-q)^2} \exp(-128np). \end{aligned}$$

For $\mathbb{E}H'_{0,i}$, note that for any positive random variable X and any $s \geq 0$, we have $\mathbb{E}X^2 \mathbb{I}\{X \geq s\} = \sum_{j=1}^{\infty} \mathbb{E}X^2 \mathbb{I}\{(j+1)s > X \geq js\} \leq \sum_{j=1}^{\infty} (j+1)^2 s^2 \mathbb{E}\mathbb{I}\{X \geq js\}$. Then

$$\begin{aligned} \mathbb{E}H'_{0,i} &\leq 25 \mathbb{E} \left(\frac{8k}{\beta n(p-q)} \|(A_{i \cdot} - \mathbb{E}A_{i \cdot})U^*\| \right)^2 \\ &\quad \times \mathbb{I} \left\{ \frac{8k}{\beta n(p-q)} \|(A_{i \cdot} - \mathbb{E}A_{i \cdot})U^*\| \geq \frac{t}{5} \right\} \\ &\leq t^2 \sum_{j=1}^{\infty} (j+1)^2 \\ &\quad \times \mathbb{P} \left(\frac{8k}{\beta n(p-q)} \|(A_{i \cdot} - \mathbb{E}A_{i \cdot})U^*\| \geq \frac{t}{5} j \right) \\ &\leq t^2 \sum_{l \in [k]} \sum_{j=1}^{\infty} (j+1)^2 \\ &\quad \times \mathbb{P} \left(|(A_{i \cdot} - \mathbb{E}A_{i \cdot})U^*_l| \geq \frac{\beta n(p-q)}{8k} \frac{t}{5k} j \right) \\ &\leq 2t^2 \sum_{l \in [k]} \sum_{j=1}^{\infty} (j+1)^2 \\ &\quad \times \exp \left(- \frac{\frac{1}{2} j^2 \left(\frac{\beta n(p-q)}{8k} \frac{t}{5k} \right)^2}{p + \frac{1}{3} j \frac{\beta n(p-q)}{8k} \frac{t}{5k} \sqrt{\frac{k}{\beta n}}} \right), \end{aligned}$$

where in the last inequality, we use Bernstein inequality and the fact that $\|U^*_l\| = 1$ and $\|U^*_l\|_{\infty} \leq \sqrt{\frac{k}{\beta n}}$ from Lemma 3.

For any $t > \sqrt{\frac{k}{\beta n}}$, we have $\frac{1}{p} \left(\frac{\beta n(p-q)}{8k} \frac{t}{5k} \right)^2 \rightarrow \infty$ and also $\left(\frac{\beta n(p-q)}{8k} \frac{t}{5k} \right) \sqrt{\frac{\beta n}{k}} \rightarrow \infty$, under the assumption that $\frac{n(p-q)^2}{\beta^{-1} k^3 p} \rightarrow \infty$. By Lemma 12, we have

$$\begin{aligned} \mathbb{E}H'_{0,i} &\leq 2 t^2 \sum_{l \in [k]} \sum_{j=1}^{\infty} (j+1)^2 \exp \left(- \frac{1}{2} j^2 \right. \\ &\quad \times \left(\frac{1}{p} \left(\frac{\beta n(p-q)}{8k} \frac{t}{5k} \right)^2 \right)^{-1} \\ &\quad \left. + \frac{1}{3} j \left(\left(\frac{\beta n(p-q)}{8k} \frac{t}{5k} \right) \sqrt{\frac{\beta n}{k}} \right)^{-1} \right)^{-1} \\ &\leq 16t^2 k \exp \left(- \frac{1}{2} \left(\frac{1}{p} \left(\frac{\beta n(p-q)}{8k} \frac{t}{5k} \right)^2 \right)^{-1} \right. \\ &\quad \left. + \frac{1}{3} \left(\left(\frac{\beta n(p-q)}{8k} \frac{t}{5k} \right) \sqrt{\frac{\beta n}{k}} \right)^{-1} \right)^{-1} \\ &= 16t^2 k \exp \left(- \frac{\frac{1}{2} \left(\frac{\beta n(p-q)}{8k} \frac{t}{5k} \right)^2}{p + \frac{1}{3} \frac{\beta n(p-q)}{8k} \frac{t}{5k} \sqrt{\frac{k}{\beta n}}} \right). \end{aligned}$$

The analysis for $\mathbb{E}H'_{3,i}$ is similar to that of $\mathbb{E}H'_{0,i}$ but is more involved. By the same argument as in the analysis of $\mathbb{E}H'_{0,i}$, we have

$$\mathbb{E}H'_{3,i}$$

$$\begin{aligned}
&\leq 25\mathbb{E}\left(\frac{8k}{\beta n(p-q)}\left\|\left(\check{A}_{i\cdot}-\check{P}_{i\cdot}\right)\left(I_n-U^*U^{*T}\right)\bar{U}^{(i)}\right\|\right)^2\times \\
&\quad \mathbb{I}\left\{\frac{8k}{\beta n(p-q)}\left\|\left(\check{A}_{i\cdot}-\check{P}_{i\cdot}\right)\left(I_n-U^*U^{*T}\right)\bar{U}^{(i)}\right\|\geq\frac{t}{5}j\right\}\mathbb{I}\{\mathcal{F}\} \\
&\leq t^2\sum_{j=1}^{\infty}(j+1)^2\mathbb{E}\mathbb{E}\left\{\frac{8k}{\beta n(p-q)}\right. \\
&\quad \left.\times\left\|\left(\check{A}_{i\cdot}-\check{P}_{i\cdot}\right)\left(I_n-U^*U^{*T}\right)\bar{U}^{(i)}\right\|\geq\frac{t}{5}j\right\}\mathbb{I}\{\mathcal{F}\}.
\end{aligned}$$

Consider a fixed $j \geq 1$ and any $t > \sqrt{\frac{k}{\beta n}}$. Follow the analysis of $\mathbb{E}H_{3,i}$ in the proof of Theorem 2, we have

$$\begin{aligned}
&\mathbb{E}\mathbb{I}\left\{\frac{8k}{\beta n(p-q)}\left\|\left(\check{A}_{i\cdot}-\check{P}_{i\cdot}\right)\left(I_n-U^*U^{*T}\right)\bar{U}^{(i)}\right\|\geq\frac{t}{5}j\right\}\mathbb{I}\{\mathcal{F}\} \\
&= \mathbb{E}\mathbb{I}\left\{\left\|\left(\check{A}_{i\cdot}-\check{P}_{i\cdot}\right)\left(I_n-U^*U^{*T}\right)\bar{U}^{(i)}\right\|\geq\frac{\beta n(p-q)tj}{40k}\right\}\mathbb{I}\{\mathcal{F}\} \\
&\leq k\sum_{\alpha\in\{-1,1\}}\mathbb{E}\left(\sup_{w\in\mathcal{W}'}\right. \\
&\quad \left.\times\mathbb{P}\left(\alpha\sum_{j\in S_i^c}(A_{ij}-\mathbb{E}A_{ij})w_j\geq\frac{\beta n(p-q)tj}{80k^2}\middle|S_i\cup\{i\}\right)\right)
\end{aligned}$$

which is analogous to (46). Here $\mathcal{W}' := \{w \in \mathbb{R}^n : \|w\| \leq \frac{4(C_0+3)k\sqrt{kn}p}{\beta n(p-q)}, \|w\|_{\infty} \leq 2t\}$, analogous to the definition of \mathcal{W} in (45). Consider any $\alpha \in \{-1, 1\}$ and any $w \in \mathcal{W}'$. Analogous to (47), we can obtain

$$\begin{aligned}
&\log \mathbb{P}\left(\alpha\sum_{j\in S_i^c}(A_{ij}-\mathbb{E}A_{ij})w_j\geq\frac{\beta n(p-q)tj}{80k^2}\middle|S_i\cup\{i\}\right) \\
&\leq -\frac{\beta n(p-q)tj}{80k^2}s + ps^2\left(\frac{4(C_0+3)k\sqrt{kn}p}{\beta n(p-q)}\right)^2\exp(2st),
\end{aligned}$$

for any $s > 0$. To derive the above display, we use Lemma 10 and assume $p \leq 1/2$. Choose $s = \frac{160k^2(p-q)}{\beta t}$ such that $\frac{\beta n(p-q)t}{80k^2}s = \frac{2n(p-q)^2}{p}$. Then we have

$$\begin{aligned}
&\log \mathbb{P}\left(\alpha\sum_{j\in S_i^c}(A_{ij}-\mathbb{E}A_{ij})w_j\geq\frac{\beta n(p-q)tj}{80k^2}\middle|S_i\cup\{i\}\right) \\
&\leq -\frac{2n(p-q)^2}{p}j + p\left(\frac{160k^2(p-q)}{\beta t}\frac{p}{p}\right)^2 \\
&\quad \times\left(\frac{4(C_0+3)k\sqrt{kn}p}{\beta n(p-q)}\right)^2\exp\left(\frac{320k^2(p-q)}{p}\right) \\
&= -\frac{2n(p-q)^2}{p}j + \left(\frac{640(C_0+3)k^{3.5}}{\beta^2\sqrt{nt}}\right)^2 \\
&\quad \times\exp\left(\frac{320k^2(p-q)}{p}\right).
\end{aligned}$$

Then

$$\begin{aligned}
&\mathbb{E}\mathbb{I}\left\{\frac{8k}{\beta n(p-q)}\left\|\left(\check{A}_{i\cdot}-\check{P}_{i\cdot}\right)\left(I_n-U^*U^{*T}\right)\bar{U}^{(i)}\right\|\geq\frac{t}{5}j\right\}\mathbb{I}\{\mathcal{F}\} \\
&\leq 2k\exp\left(-\frac{2n(p-q)^2}{p}j\right)
\end{aligned}$$

and consequently,

$$\begin{aligned}
\mathbb{E}H'_{3,i} &\leq 2k^2t^2\sum_{j=1}^{\infty}(j+1)^2\exp\left(-\frac{2n(p-q)^2}{p}j\right. \\
&\quad \left.+\left(\frac{640(C_0+3)k^{3.5}}{\beta^2\sqrt{nt}}\right)^2\exp\left(\frac{320k^2(p-q)}{p}\right)\right) \\
&\leq 16k^2t^2\exp\left(-\frac{2n(p-q)^2}{p}\right. \\
&\quad \left.+\left(\frac{640(C_0+3)k^{3.5}}{\beta^2\sqrt{nt}}\right)^2\exp\left(\frac{320k^2(p-q)}{p}\right)\right),
\end{aligned}$$

where the last inequality is due to that $\frac{n(p-q)^2}{p} \rightarrow \infty$.

Combining all the above expressions together, (54) leads to

$$\begin{aligned}
&\mathbb{E}\sum_{i\in[n]}\|U_{i\cdot}\|^2\mathbb{I}\{\|U_{i\cdot}\|\geq t\} \\
&\leq 4\frac{40^2k^3p^2}{\beta^3n(p-q)^2}n\exp(-128np) \\
&\quad + 12\frac{160k^2t^2np^2}{\beta^2n(p-q)^2}n\exp(-128np) \\
&\quad + 32t^2kn\exp\left(-\frac{\frac{1}{2}\left(\frac{\beta n(p-q)}{8k}\frac{t}{5k}\right)^2}{p+\frac{1}{3}\frac{\beta n(p-q)}{8k}\frac{t}{5k}\sqrt{\frac{k}{\beta n}}}\right) \\
&\quad + 32k^2t^2n\exp\left(-\frac{2n(p-q)^2}{p}\right. \\
&\quad \left.+\left(\frac{640(C_0+3)k^{3.5}}{\beta^2\sqrt{nt}}\right)^2\exp\left(\frac{320k^2(p-q)}{p}\right)\right),
\end{aligned}$$

for any $t > 20\sqrt{\frac{k}{\beta n}}$. Taking $t = t_0$ where t_0 is defined in (13), we have

$$\begin{aligned}
&\mathbb{E}\sum_{i\in[n]}\|U_{i\cdot}\|^2\mathbb{I}\{\|U_{i\cdot}\|\geq t_0\} \\
&\leq 4\frac{40^2p}{k^2n(p-q)^2}nt_0^2(np)\exp(-128np) \\
&\quad + 12\frac{160k^2p}{\beta^2n(p-q)^2}nt_0^2(np)\exp(-128np) \\
&\quad + 32t_0^2kn\exp\left(-\frac{2n(p-q)^2}{p}\right) \\
&\quad + 32k^2t_0^2n\exp\left(-\frac{2n(p-q)^2}{p}\right. \\
&\quad \left.+ 72k^2(C_0+3)^2\beta^{-1}\exp\left(\frac{320k^2(p-q)}{p}\right)\right) \\
&\leq t_0^2n\exp\left(-\frac{3n(p-q)^2}{2p}\right),
\end{aligned}$$

where the last inequality holds under the assumption that $\beta^{-1}, k = O(1)$, $0 < q < p \leq 1/2$, and $\frac{n(p-q)^2}{p} \rightarrow \infty$.

APPENDIX A
PROOF OF THEOREM 4

The proof of Theorem 4 follows the proof of Theorem 2 with some modifications as we need to derive a lower bound instead of an upper bound. Define $(a, b) := \operatorname{argmin}_{1 \leq a' \neq b' \leq k} J_{n_{a'}, n_{b'}, p, q}$. We have

$$\begin{aligned} \ell(\hat{z}, z^*) &\geq \ell(\hat{z}, z^*) \mathbb{I}\{\mathcal{F}\} \geq \frac{1}{n} \sum_{i \notin S} \mathbb{I}\{\hat{z}_i \neq z_i^*\} \mathbb{I}\{\mathcal{F}\} \\ &\geq \frac{1}{n} \sum_{i \notin S: z_i^* = b} \mathbb{I}\{\hat{z}_i \neq z_i^*\} \mathbb{I}\{\mathcal{F}\} \\ &\geq \frac{1}{n} \sum_{i \notin S: z_i^* = b} \mathbb{I}\left\{\left\|U_i \Lambda - \hat{\theta}_a\right\| \leq \left\|U_i \Lambda - \hat{\theta}_{z_i^*}\right\|\right\} \mathbb{I}\{\mathcal{F}\}, \end{aligned}$$

where in the second inequality we use (15). Consider any $i \notin S$ such that $z_i^* = b$. Analogous to the establishment of (38), we can have

$$\begin{aligned} &\mathbb{I}\left\{\left\|U_i \Lambda - \hat{\theta}_a\right\| \leq \left\|U_i \Lambda - \hat{\theta}_{z_i^*}\right\|\right\} \mathbb{I}\{\mathcal{F}\} \\ &\geq \mathbb{I}\left\{(p-q) \left(\sum_{j: z_j^* = a} (A_{ij} - \mathbb{E}A_{ij}) - \sum_{j: z_j^* = z_i^*, j \neq i} (A_{ij} - \mathbb{E}A_{ij})\right) \right. \\ &\quad \left. - 2C_2 \beta^{-0.5} k \sqrt{p} \|(A_i - \mathbb{E}A_i)U^*\| \right. \\ &\quad \left. - 2\sqrt{n_a + n_{z_i^*}}(p-q) \|(A_i - \mathbb{E}A_i)(I_n - U^*U^{*T})U\| \right\} \mathbb{I}\{\mathcal{F}\} \\ &\geq \frac{1}{2} \left(1 - C_3 \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}}\right) (n_a + n_{z_i^*})(p-q)^2 \mathbb{I}\{\mathcal{F}\}. \end{aligned}$$

Consider a positive sequence $\rho = o(1)$ whose value will be determined later. Analogous to the establishment of (39), we have

$$\begin{aligned} &\mathbb{I}\left\{\left\|U_i \Lambda - \hat{\theta}_a\right\| \leq \left\|U_i \Lambda - \hat{\theta}_{z_i^*}\right\|\right\} \mathbb{I}\{\mathcal{F}\} \\ &\geq \mathbb{I}\left\{\sum_{j: z_j^* = a} (A_{ij} - \mathbb{E}A_{ij}) - \sum_{j: z_j^* = z_i^*, j \neq i} (A_{ij} - \mathbb{E}A_{ij}) \geq \frac{1}{2} \times \right. \\ &\quad \left(1 + 4\rho + C_3 \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}}\right) (n_a + n_{z_i^*})(p-q) \Big\} \mathbb{I}\{\mathcal{F}\} \\ &\quad - \mathbb{I}\left\{\|(A_i - \mathbb{E}A_i)U^*\| \geq \frac{\rho n(p-q)^2}{C_2 \beta^{-1.5} k^2 \sqrt{p}}\right\} \\ &\quad - \mathbb{I}\left\{\|(A_i - \mathbb{E}A_i)(I_n - U^*U^{*T})U\| \geq \right. \\ &\quad \left. \rho \sqrt{\frac{\beta n}{k}}(p-q)\right\} \mathbb{I}\{\mathcal{F}\} \\ &=: G'_{1,i,a} \mathbb{I}\{\mathcal{F}\} - G_{2,i} - G_{3,i}. \end{aligned}$$

Note here $G'_{1,i,a}$ is slightly different from $G_{1,i,a}$ in the proof of Theorem 2 and $G_{2,i}, G_{3,i}$ are exactly the same quantities as they appear in the proof of Theorem 2. Then the decomposition (41) also holds for $G_{3,i}$. Then analogous to (42), we have

$$\begin{aligned} \mathbb{E}\ell(\hat{z}, z^*) &\geq \frac{1}{n} \mathbb{E} \sum_{i \notin S: z_i^* = b} (G'_{1,i,a} \mathbb{I}\{\mathcal{F}\} - G_{2,i} - G_{3,i}) \\ &\geq \frac{1}{n} \mathbb{E} \sum_{i \notin S: z_i^* = b} G'_{1,i,a} \mathbb{I}\{\mathcal{F}\} - \frac{1}{n} \sum_{i \in [n]} \mathbb{E}G_{2,i} \end{aligned}$$

$$\begin{aligned} &-\frac{1}{n} \sum_{i \in [n]} \mathbb{E}H_{1,i} - \frac{1}{n} \sum_{i \in [n]} \mathbb{E}H_{2,i} \\ &-\frac{1}{n} \sum_{i \in [n]} (\mathbb{E}H_{3,i} + \mathbb{E}H_{4,i} + \mathbb{E}H_{5,i}). \end{aligned}$$

In the above display, we only need to analyze $\mathbb{E} \sum_{i \notin S: z_i^* = b} G'_{1,i,a} \mathbb{I}\{\mathcal{F}\}$ as all the remaining terms have been analyzed in the proof of Theorem 2. Consider any $i \in [n]$ and any $z_i^* = b$. We introduce an auxiliary random variable $X \sim \text{Bernoulli}(p)$ that is independent of A . Since $X \geq 0$, we have

$$\begin{aligned} &\mathbb{E}G'_{1,i,a} \\ &= \mathbb{P}\left(\sum_{j: z_j^* = a} (A_{ij} - \mathbb{E}A_{ij}) - \sum_{j: z_j^* = z_i^*, j \neq i} (A_{ij} - \mathbb{E}A_{ij}) \right. \\ &\quad \left. \geq \frac{1}{2} \left(1 + 4\rho + C_3 \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}}\right) (n_a + n_{z_i^*})(p-q) \right) \\ &\geq \mathbb{P}\left(\sum_{j: z_j^* = a} (A_{ij} - \mathbb{E}A_{ij}) - \sum_{j: z_j^* = z_i^*, j \neq i} (A_{ij} - \mathbb{E}A_{ij}) \geq \frac{1}{2} \times \right. \\ &\quad \left. \left(1 + 4\rho + C_3 \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}}\right) (n_a + n_{z_i^*})(p-q) + X \right) \\ &= \mathbb{P}\left(\sum_{j: z_j^* = a} (A_{ij} - \mathbb{E}A_{ij}) - \sum_{j: z_j^* = z_i^*, j \neq i} (A_{ij} - \mathbb{E}A_{ij}) \right. \\ &\quad \left. - (X - p) \geq \frac{1}{2} \left(1 + 4\rho + C_3 \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}}\right) \right. \\ &\quad \left. \times (n_a + n_{z_i^*})(p-q) + p \right) \\ &\geq \mathbb{P}\left(\sum_{j: z_j^* = a} (A_{ij} - \mathbb{E}A_{ij}) - \sum_{j: z_j^* = z_i^*, j \neq i} (A_{ij} - \mathbb{E}A_{ij}) \right. \\ &\quad \left. - (X - p) \geq \frac{1}{2} \left(1 + 4\rho + C_3 \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}} + \frac{pk}{\beta n(p-q)}\right) \right. \\ &\quad \left. \times (n_a + n_{z_i^*})(p-q) \right) \\ &\geq \exp\left(-J_{n_a, n_b, p, q} - \left(4\rho + C_3 \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}} + \frac{pk}{\beta n(p-q)}\right) \right. \\ &\quad \left. \times (n_a + n_{z_i^*}) \frac{(p-q)^2}{p} - 4\sqrt{\frac{(n_a + n_{z_i^*})(p-q)^2 p}{q^2}} \right) \\ &\quad \times \left(\frac{1}{4} - \sqrt{\frac{2}{(n_a + n_{z_i^*})q}}\right). \end{aligned}$$

where we use Lemma 15 in the last inequality. Under the assumption $\frac{n(p-q)^2}{\beta^{-1}kp} \rightarrow \infty$ and $p/q = O(1)$, using Lemma 13, we have

$$\begin{aligned} &\mathbb{E}G'_{1,i,a} \\ &\geq \frac{1}{8} \exp\left(-\left(1 + \left(4\rho + C_3 \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}} + \frac{pk}{\beta n(p-q)}\right) \right. \right. \end{aligned}$$

$$\times \frac{16kp}{\beta q} + 64 \frac{kp}{\beta q} \sqrt{\frac{p}{n(p-q)^2}} \Big) J_{n_a, n_b, p, q} \Big).$$

Note that $J_{n_a, n_b, q, q} = \min_{1 \leq a' \neq b' \leq k} J_{n_{a'}, n_{b'}, p, q} = J_{\min}$. We have

$$\begin{aligned} & \mathbb{E} \sum_{i \notin S: z_i^* = b} G'_{1, i, a} \mathbb{I}\{\mathcal{F}\} \\ & \geq \mathbb{E} \sum_{i \in [n]: z_i^* = b} G'_{1, i, a} - |S| - n\mathbb{P}(\mathcal{F}^c) \\ & \geq \frac{\beta n}{8k} \exp \left(- \left(1 + \left(4\rho + C_3 \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}} + \frac{pk}{\beta n(p-q)} \right) \right. \right. \\ & \quad \times \frac{16kp}{\beta q} + 64 \frac{kp}{\beta q} \sqrt{\frac{p}{n(p-q)^2}} \Big) J_{\min} \Big) \\ & \quad - n \exp(-128np) - 2n^{-2}, \end{aligned}$$

where we use Lemma 6. Analogous to the establishment of (48), we have

$$\begin{aligned} & \mathbb{E} \ell(\hat{z}, z^*) \\ & \geq \frac{\beta}{8k} \exp \left(- \left(1 + \left(4\rho + C_3 \sqrt{\frac{\beta^{-2} k^3 p}{n(p-q)^2}} + \frac{pk}{\beta n(p-q)} \right) \right. \right. \\ & \quad \times \frac{16kp}{\beta q} + 64 \frac{kp}{\beta q} \sqrt{\frac{p}{n(p-q)^2}} \Big) J_{\min} \Big) \\ & \quad - \exp(-128np) - 2n^{-3} - 4k^2 \exp \left(- \frac{3n(p-q)^2}{2p} \right), \end{aligned}$$

under the assumption $\beta^{-1}, k = O(1), 1 \leq q < p \leq 1/2, p/q = O(1), \frac{n(p-q)^2}{p} \rightarrow \infty$, and $\rho^{-1} = o\left(\log\left(\frac{n(p-q)^2}{p}\right)\right)$. We can take $\rho^{-1} = \left(\log\left(\frac{n(p-q)^2}{p}\right)\right)^{\frac{1}{4}}$. When $p \leq 1/10$ is additionally assume, by following same argument as in the proof of Theorem 2, there exists some constant C_5 such that

$$\begin{aligned} & \mathbb{E} \ell(\hat{z}, z^*) \geq \\ & \exp \left(- \left(1 + C_5 \left(\log \left(\frac{n(p-q)^2}{p} \right) \right)^{-\frac{1}{4}} \right) J_{\min} \right) - 2n^{-3}. \end{aligned}$$

APPENDIX B PROOFS OF RESULTS IN SECTION II

First we give an equivalent expression for P . Define $Z^* \in \{0, 1\}^{n \times k}$ to be a matrix such that $Z_{ij}^* = \mathbb{I}\{z_j^* = i\}$ for all $i \in [n]$ and $j \in [k]$. In addition, define $B \in \mathbb{R}^{k \times k}$ such that $B_{ab} = q\mathbb{I}\{a \neq b\} + p\mathbb{I}\{a = b\}$ for all $a, b \in [k]$. Then we can verify

$$P = Z^* B Z^{*T}.$$

Then following lemmas are about properties of population quantities.

Lemma 3: Define $\Delta := \text{diag}(\sqrt{n_1}, \dots, \sqrt{n_k})$. There exists some $W \in \mathcal{O}(k, k)$ such that $U^* = Z^* \Delta^{-1} W$ and $\Lambda^* = W^T \Delta B \Delta W$. In addition,

$$\|U_{i \cdot}^*\| = \frac{1}{\sqrt{n_{z_i^*}}}, \forall i \in [n].$$

Consequently,

$$\max_{i \in [n], j \in [k]} \|U_{ij}^*\| \leq \|U^*\|_{2, \infty} = \sqrt{k/(\beta n)}.$$

Proof: Note that $P = Z^* B Z^{*T} = Z^* \Delta^{-1} (\Delta B \Delta) \Delta^{-1} Z^{*T}$ and $Z^* \Delta^{-1} \in \mathcal{O}(k, k)$. Apply SVD to the matrix $\Delta B \Delta$ and we obtain $\Delta B \Delta = W \tilde{\Lambda} W^T$ for some $W \in \mathcal{O}(k, k)$ and some diagonal matrix $\tilde{\Lambda}$. Since $Z^* \Delta^{-1} W \in \mathcal{O}(n, k)$ and $P = (Z^* \Delta^{-1} W) \tilde{\Lambda} (Z^* \Delta^{-1} W)^T$, we have $\Lambda^* = \tilde{\Lambda} = W^T \Delta B \Delta W$ and $U = Z^* \Delta^{-1} W$. As a result, for each $i \in [n]$, we have

$$\|U_{i \cdot}\| = \|Z_{i \cdot}^* \Delta^{-1} W\| = \frac{1}{\sqrt{n_{z_i^*}}} \|W_{z_i^* \cdot}\| = \frac{1}{\sqrt{n_{z_i^*}}}.$$

□

Lemma 4: For any $a, b \in [k]$ such that $a \neq b$, we have $U^*(\theta_a^* - \theta_b^*)^T \in \mathbb{R}^n$ satisfies

$$(U^*(\theta_a^* - \theta_b^*)^T)_j = \begin{cases} (p-q), & \text{if } z_j^* = a, \\ -(p-q), & \text{if } z_j^* = b, \\ 0, & \text{o.w.,} \end{cases} \quad (55)$$

for all $j \in [n]$.

Proof: Let $i, i' \in [n]$ be any two indexes such that $z_i^* = a$ and $z_{i'}^* = b$. From Lemma 1, we have

$$\begin{aligned} U^*(\theta_a^* - \theta_b^*)^T &= U^*((U^* \Lambda^*)_{i \cdot} - (U^* \Lambda^*)_{i' \cdot})^T \\ &= U^*(U_{i \cdot}^* \Lambda^* - U_{i' \cdot}^* \Lambda^*)^T \\ &= U^* \Lambda^* (U_{i \cdot}^* - U_{i' \cdot}^*)^T = P_{i \cdot} - P_{i' \cdot}. \end{aligned}$$

The proof is complete with $P_{ji} = p\mathbb{I}\{z_j^* = a\} + q\mathbb{I}\{z_j^* \neq a\}$ and $P_{ji'} = p\mathbb{I}\{z_j^* = b\} + q\mathbb{I}\{z_j^* \neq b\}$ for any $j \in [n]$. □

Proof of Lemma 1: From Lemma 3, we have $U^* = Z^* \Delta^{-1} W$ and $\Lambda^* = W^T \Delta B \Delta W$ for some $W \in \mathcal{O}(k, k)$. Then

$$U^* \Lambda^* = (Z^* \Delta^{-1} W)(W^T \Delta B \Delta W) = Z^* B \Delta W.$$

Since Z^* has k unique rows, so does $U^* \Lambda^*$. Denote $\theta_a^* = B_{a \cdot} \Delta W$ for each $a \in [k]$. Then

$$\begin{aligned} \|\theta_a^*\|^2 &= \|B_{a \cdot} \Delta\|^2 = \sum_{b \in [k]} (B_{ab} \Delta_{bb})^2 = \sum_{b \in [k]} (B_{ab} n_b)^2 \\ &= (p^2 - q^2) n_a + q^2 n, \forall a \in [k], \end{aligned}$$

where we use the fact that $\sum_{b \in [k]} n_b = n$. For each $i \in [n]$, we have $(U^* \Lambda^*)_{i \cdot} = Z_{i \cdot}^* B \Delta W = \theta_{z_i^*}^*$. We also have for any $a, b \in [k]$ such that $a \neq b$.

$$\begin{aligned} \|\theta_a^* - \theta_b^*\| &= \|(B_{a \cdot} - B_{b \cdot}) \Delta W\| = \|(B_{a \cdot} - B_{b \cdot}) \Delta\| \\ &= (p-q) \sqrt{n_a + n_b}. \end{aligned}$$

□

Proof of Proposition 1: Note that performing k -means on $\{U_{i \cdot} \Lambda\}_{i \in [n]}$ is equivalent to performing k -means on $\{U_{i \cdot} \Lambda U^T\}_{i \in [n]}$. This is because U has orthonormal columns and consequently $\|U_{i \cdot} \Lambda - U_{j \cdot} \Lambda\| = \|U_{i \cdot} \Lambda U^T - U_{j \cdot} \Lambda U^T\|$ for all $i, j \in [n]$. Since $U_{i \cdot} \Lambda U^T = \hat{A}_{i \cdot} U U^T$, we have

$$\begin{aligned} & (\hat{z}, \{\hat{\theta}_1, \dots, \hat{\theta}_k\}) \\ &= \underset{z \in [k]^n, \theta_1, \dots, \theta_k \in \mathbb{R}^{1 \times k}}{\text{argmin}} \sum_{i \in [n]} \|U_{i \cdot} \Lambda U^T - \theta_{z_i} U^T\|^2 \end{aligned}$$

$$= \underset{z \in [k]^n, \theta_1, \dots, \theta_k \in \mathbb{R}^{1 \times k}}{\operatorname{argmin}} \sum_{i \in [n]} \left\| \tilde{A}_i U U^T - \theta_{z_i}^* U^T \right\|^2.$$

Note that P has k unique rows $\theta_1^* U^{*T}, \dots, \theta_k^* U^{*T}$. According to Lemma 1,

$$\begin{aligned} \min_{a, b \in [k]: a \neq b} \left\| \theta_a^* U^{*T} - \theta_b^* U^{*T} \right\| &= \min_{a, b \in [k]: a \neq b} \left\| \theta_a^* - \theta_b^* \right\| \\ &= \min_{a, b \in [k]: a \neq b} \sqrt{n_a + n_b} (p - q) =: \delta. \end{aligned}$$

Hence, δ is the minimum distance among all k unique rows of P . Then by Proposition 3.1 of [33], if

$$\frac{\delta}{\beta^{-0.5} k n^{-0.5} \left\| \tilde{A} - P \right\|} \geq 16 \quad (56)$$

is satisfied, there exists some $\phi \in \Phi$ and some constant $C > 0$ such that

$$\frac{1}{n} \sum_{i \in [n]} \mathbb{I} \{ \hat{z}_i \neq \phi(z_i^*) \} \leq \frac{Ck \left\| \tilde{A} - P \right\|^2}{n\delta^2},$$

and,

$$\max_{a \in [k]} \left\| \hat{\theta}_{\phi(a)} U^T - \theta_a^* U^{*T} \right\| \leq C\beta^{-0.5} k n^{-0.5} \left\| \tilde{A} - P \right\|.$$

In the following, we are going to give an upper bound for $\left\| \tilde{A} - P \right\|$. Note that $\left\| \tilde{A} - P \right\| \leq \left\| \tilde{A} - \mathbb{E}A \right\| + \left\| \mathbb{E}A - P \right\| = \left\| \tilde{A} - \mathbb{E}A \right\| + p$ with $\left\| \tilde{A} - \mathbb{E}A \right\| \leq C_0 \sqrt{np}$ assumed. In addition, we have $\delta \geq \sqrt{2\beta n/k} (p - q)$. As a result, there exists a constant $C_1 > 0$, such that if $\frac{n(p-q)^2}{\beta^{-2} k^3 p} \geq C_1'$, we have (56) satisfied, and consequently obtain the desired upper bounds in (5) and (6) for some constant $C_2 > 0$. \square

Proof of Proposition 2: First consider any $i \in [n]$ such that $z_i^* = 2$. Then

$$\begin{aligned} \mathbb{I} \{ \tilde{z}_i \neq z_i^* \} &= \mathbb{I} \left\{ \left\| A_i U^* - \theta_1^* \right\|^2 \leq \left\| A_i U^* - \theta_2^* \right\|^2 \right\} \\ &= \mathbb{I} \left\{ 2A_i U^* (\theta_1^* - \theta_2^*)^T \geq -\left\| \theta_2^* \right\|^2 + \left\| \theta_1^* \right\|^2 \right\} \\ &= \mathbb{I} \left\{ 2(p - q) \left(\sum_{j: z_j^* = 1} A_{ij} - \sum_{j \neq i: z_j^* = 2} A_{ij} \right) \right. \\ &\quad \left. \geq (p^2 - q^2) (n_1 - n_2) \right\} \\ &= \mathbb{I} \left\{ \sum_{j: z_j^* = 1} A_{ij} - \sum_{j \neq i: z_j^* = 2} A_{ij} \geq \frac{p+q}{2} (n_1 - n_2) \right\} \\ &= \mathbb{I} \left\{ \sum_{j: z_j^* = 1} (A_{ij} - \mathbb{E}A_{ij}) - \sum_{j \neq i: z_j^* = 2} (A_{ij} - \mathbb{E}A_{ij}) \right. \\ &\quad \left. \geq \frac{p-q}{2} n - p \right\}, \quad (57) \end{aligned}$$

where the third equation is by Lemma 1 and Lemma 4 and in the last equation we use $n_1 + n_2 = n$. Note that $|\{j \neq i : z_j^* = 2\}| = n_2 - 1$. Using Lemma 15, we have

$$\mathbb{E} \mathbb{I} \{ \tilde{z}_i \neq z_i^* \}$$

$$\begin{aligned} &= \mathbb{P} \left(\sum_{j: z_j^* = 1} (A_{ij} - \mathbb{E}A_{ij}) - \sum_{j \neq i: z_j^* = 2} (A_{ij} - \mathbb{E}A_{ij}) \right. \\ &\quad \left. \geq \frac{1}{2} \left(1 - \frac{2q}{(n-1)(p-q)} \right) (n-1)(p-q) \right) \\ &\leq \exp \left(-J_{n_1, n_2-1, p, q} \right. \\ &\quad \left. + \frac{q}{(n-1)(p-q)} (n_1 + n_2 - 1) \frac{(p-q)^2}{q} \right) \\ &= \exp \left(-J_{n_1, n_2-1, p, q} + (p-q) \right). \end{aligned}$$

From Lemma 13, we have

$$\begin{aligned} \mathbb{E} \mathbb{I} \{ \tilde{z}_i \neq z_i^* \} &\leq \exp \left(-J_{n_1, n_2, p, q} + \frac{(p-q)^2}{4q} + (p-q) \right) \\ &\leq \exp \left(- \left(1 - \frac{(p-q)^2}{4q} + (p-q) \right) J_{n_1, n_2, p, q} \right) \\ &\leq \exp \left(- \left(1 - \frac{(p-q)^2}{4q} + (p-q) \right) J_{n_1, n_2, p, q} \right) \\ &\leq \exp \left(- \left(1 - \frac{C_1}{n} \right) J_{n_1, n_2, p, q} \right), \end{aligned}$$

for some constant $C_1 > 0$. For its lower bound, we introduce an auxiliary random variable $X \sim \text{Bernoulli}(p)$ that is independent of A . Since $X \geq 0$, (57) leads to

$$\begin{aligned} \mathbb{I} \{ \tilde{z}_i \neq z_i^* \} &\geq \mathbb{I} \left\{ \sum_{j: z_j^* = 1} (A_{ij} - \mathbb{E}A_{ij}) - \sum_{j \neq i: z_j^* = 2} (A_{ij} - \mathbb{E}A_{ij}) \right. \\ &\quad \left. \geq \frac{p-q}{2} n + X - p \right\} \\ &= \mathbb{I} \left\{ \sum_{j: z_j^* = 1} (A_{ij} - \mathbb{E}A_{ij}) - \sum_{j \neq i: z_j^* = 2} (A_{ij} - \mathbb{E}A_{ij}) \right. \\ &\quad \left. - (X - p) \geq \frac{p-q}{2} n \right\}. \end{aligned}$$

Using Lemma 15 and Lemma 13 again, we have

$$\begin{aligned} \mathbb{E} \mathbb{I} \{ \tilde{z}_i \neq z_i^* \} &\geq \exp \left(-J_{n_1, n_2, p, q} - 4\sqrt{\frac{n(p-q)^2 p}{q^2}} \right) \left(\frac{1}{4} - \sqrt{\frac{2}{nq}} \right) \\ &\geq \exp \left(- \left(1 + \frac{4\sqrt{\frac{n(p-q)^2 p}{q^2}}}{J_{n_1, n_2, p, q}} \right) J_{n_1, n_2, p, q} \right) \left(\frac{1}{4} - \sqrt{\frac{2}{nq}} \right) \\ &\geq \exp \left(- \left(1 + \frac{4\sqrt{\frac{n(p-q)^2 p}{q^2}}}{n_2 \frac{(p-q)^2}{8p}} \right) J_{n_1, n_2, p, q} \right) \left(\frac{1}{4} - \sqrt{\frac{2}{nq}} \right) \\ &\geq \exp \left(- \left(1 + C_2 \sqrt{\frac{p}{n(p-q)^2}} \right) J_{n_1, n_2, p, q} \right), \end{aligned}$$

for some constant $C_2 > 0$.

Similarly, for any $i \in [n]$ such that $z_i^* = 1$, we have

$$\begin{aligned} & \exp \left(- \left(1 + C_2 \sqrt{\frac{p}{n(p-q)^2}} \right) J_{n_2, n_1, p, q} \right) \\ & \leq \mathbb{E} \mathbb{I} \{ \tilde{z}_i \neq z_i^* \} \leq \exp \left(- \left(1 - \frac{C_1}{n} \right) J_{n_2, n_1, p, q} \right). \end{aligned}$$

Hence,

$$\begin{aligned} & \mathbb{E} \left(\frac{1}{n} \sum_{i \in [n]} \mathbb{I} \{ \tilde{z}_i \neq z_i^* \} \right) \\ & \leq \frac{1}{n} \left(n_1 \exp \left(- \left(1 - \frac{C_1}{n} \right) J_{n_2, n_1, p, q} \right) \right. \\ & \quad \left. + n_2 \exp \left(- \left(1 - \frac{C_1}{n} \right) J_{n_1, n_2, p, q} \right) \right) \\ & \leq \exp \left(- \left(1 - \frac{C_1}{n} \right) (J_{n_1, n_2, p, q} \wedge J_{n_2, n_1, p, q}) \right), \end{aligned}$$

and similarly

$$\begin{aligned} & \mathbb{E} \left(\frac{1}{n} \sum_{i \in [n]} \mathbb{I} \{ \tilde{z}_i \neq z_i^* \} \right) \\ & \geq \frac{n_1 \wedge n_2}{n} \exp \left(- \left(1 + C_2 \sqrt{\frac{p}{n(p-q)^2}} \right) \right. \\ & \quad \left. \times (J_{n_1, n_2, p, q} \wedge J_{n_2, n_1, p, q}) \right). \end{aligned}$$

APPENDIX C AUXILIARY LEMMAS

Lemma 5: We have $\lambda_k^* \geq \frac{\beta n(p-q)}{k}$ and

$$\max_{i \in [k]} |\lambda_i - \lambda_i^*|, \max_{i > k} |\lambda_i| \leq \|\tilde{A} - P\|.$$

Under the assumption that $\|\tilde{A} - P\| \leq \frac{\beta n(p-q)}{2k}$, we have

$$\|\Lambda^{-1}\| \leq \frac{2k}{\beta n(p-q)}.$$

Proof: From Lemma 3, we know $\lambda_1^*, \dots, \lambda_k^*$ are the eigenvalues of $\Delta B \Delta$. Then

$$\begin{aligned} \lambda_k^* & \geq \min_{v \in \mathbb{R}^k: \|v\|=1} v^T \Delta B \Delta v = \min_{v \in \mathbb{R}^k: \|v\|=1} (\Delta v)^T B (\Delta v) \\ & \geq (\min_{a \in [k]} n_a) \min_{v \in \mathbb{R}^k: \|v\|=1} v^T B v \\ & = (\min_{a \in [k]} n_a) \min_{v \in \mathbb{R}^k: \|v\|=1} \left(q(v^T \mathbf{1}_k)^2 + (p-q) \|v\|^2 \right) \\ & \geq (p-q) (\min_{a \in [k]} n_a) = \frac{\beta n(p-q)}{k}. \end{aligned}$$

The upper bound for $\max_{i \in [k]} |\lambda_i - \lambda_i^*|, \max_{i > k} |\lambda_i|$ is from Weyl's inequality. If $\|\tilde{A} - P\| \leq \frac{\beta n(p-q)}{2k}$ is further assumed, we have $\lambda_k \geq \frac{\beta n(p-q)}{2k}$. The proof is complete with $\|\Lambda^{-1}\| = \lambda_k^{-1}$. \square

Lemma 6: We have

$$\mathbb{E} |S| \leq n \exp(-128np).$$

Recall the definition of S_i in (23). We also have $\mathbb{E} |S_i| \leq \exp(-128np)$ for each $i \in [n]$.

Proof: We have $\mathbb{E} |S| = n\mathbb{P}(\text{Binomial}(n-1, p) \geq \tau) \leq n \exp(-128np)$ by Chernoff bound. The same upper bound holds for $\mathbb{E} |S_i|$. \square

Lemma 7: Under the assumption $\max\{p\sqrt{n}, \|\tilde{A} - P\|\} \leq \beta n(p-q)/(8k^2)$, we have

$$\inf_{W \in \mathcal{O}(k, k)} \|U - U^{(i)} W\|_F \leq 6k^{1.5} \|U_{\cdot i}\|, \forall i \in [n].$$

Proof of Lemma 7: Consider any $i \in [n]$. Note that U (resp. $U^{(i)}$) is the leading eigenspace of \tilde{A} (resp. $\tilde{A}^{(i)}$) and $\tilde{A}^{(i)}$ is obtained from \tilde{A} by zeroing out its i th row and column. In addition, from Lemma 5, we have $\lambda_k^* \geq \beta n(p-q)/k$. We also have $\max_{j \in [n]} \|P_{j\cdot}\| \leq p\sqrt{n}$.

We are going to have a partition for the eigenspaces. But before that, we need to have a partition of $[k]$. With $\lambda_0^* := \infty, \lambda_{k+1}^* := 0$, and $a_1 := 1$, define $b_1 := \arg\max\{a_1 \leq j \leq k : \lambda_{a_1}^*/(\lambda_j^* - \lambda_{j+1}^*) \leq 2k\}$. Note that the set $\{a_1 \leq j \leq k : \lambda_{a_1}^*/(\lambda_j^* - \lambda_{j+1}^*) \leq 2k\}$ is not an empty set. Otherwise, we have $\lambda_j^* - \lambda_{j+1}^* \leq \lambda_{a_1}^*/(2k)$ for all $a_1 \leq j \leq k$. Then we have $\lambda_{a_1}^* - \lambda_{k+1}^* = \sum_{j=a_1}^k (\lambda_j^* - \lambda_{j+1}^*) \leq \lambda_{a_1}^*/2$ which gives $\lambda_{k+1}^* \geq \lambda_{a_1}^*/2 \geq \lambda_k^*/2 > 0$, a contradiction with the fact that $\lambda_{k+1}^* = 0$. As a result, the aforementioned set is non-empty and b_1 is well-defined. If $b_1 < k$, then we define $a_2 := b_1 + 1$ and $b_2 := \arg\max\{a_2 \leq j \leq k : \lambda_{a_2}^*/(\lambda_j^* - \lambda_{j+1}^*) \leq 2k\}$. By the same argument as above, b_2 is also well-defined. If $b_2 < k$, we repeat this procedure until we have $b_r = k$ for some $r \leq k$. In this way, we have a partition of $[k] = \cup_{s=1}^r \{j \in \mathbb{N} : a_s \leq j \leq b_s\}$. For any $s \in [r]$, define $U_s^* := (u_{a_s}^*, \dots, u_{b_s}^*)$, $U_s := (u_{a_s}, \dots, u_{b_s})$, and define $U_s^{(i)}$ analogously. Then

$$\begin{aligned} \|UU^T - U^{(i)}U^{(i)T}\|_F^2 & = \left\| \sum_{s \in [r]} (U_s U_s^T - U_s^{(i)} U_s^{(i)T}) \right\|_F^2 \\ & \leq r \sum_{s \in [r]} \|U_s U_s^T - U_s^{(i)} U_s^{(i)T}\|_F^2. \end{aligned}$$

Now consider any $s \in [r]$. With $b_0 := 0$ and $a_{r+1} := k+1$, define $\Delta_s^* := (\lambda_{b_{s-1}}^* - \lambda_{a_s}^*) \wedge (\lambda_{b_s}^* - \lambda_{a_{s+1}}^*)$ to be the spectral gap for U_s^* . We have

$$\begin{aligned} \frac{\lambda_{a_s}^*}{\Delta_s^*} & = \frac{\lambda_{a_s}^*}{(\lambda_{b_{s-1}}^* - \lambda_{a_s}^*) \wedge (\lambda_{b_s}^* - \lambda_{a_{s+1}}^*)} \\ & = \frac{\lambda_{a_s}^*}{\lambda_{b_{s-1}}^* - \lambda_{a_s}^*} \vee \frac{\lambda_{a_s}^*}{\lambda_{b_s}^* - \lambda_{a_{s+1}}^*} \\ & \leq \frac{\lambda_{a_s}^*}{(2k)^{-1} \lambda_{a_{s-1}}^* \mathbb{I}\{s \geq 2\} + \infty \mathbb{I}\{s = 1\}} \vee (2k) = 2k. \end{aligned}$$

This implies that $\Delta_s^* \geq \lambda_{a_s}^*/(2k) \geq \lambda_k^*/2k \geq \beta n(p-q)/(2k^2)$. Under the assumption that $\max\{p\sqrt{n}, \|\tilde{A} - P\|\} \leq \beta n(p-q)/(8k^2)$, by Lemma 3 of [24], we have

$$\begin{aligned} \|U_s U_s^T - U_s^{(i)} U_s^{(i)T}\| & \leq 3(2k) \|(U_s U_s^T U_s^*)_{\cdot i}\| \\ & = 6k \|(U_s)_{\cdot i} U_s^T U_s^*\| \\ & \leq 6k \|(U_s)_{\cdot i}\|. \end{aligned}$$

Hence,

$$\begin{aligned} \|UU^T - U^{(i)}U^{(i)T}\|_F^2 &\leq \sum_{s \in [r]} 36k^2 r \|(U_s)_i\|^2 \\ &= 36k^2 r \|U_{i\cdot}\|^2 \\ &\leq 36k^3 \|U_{i\cdot}\|^2, \end{aligned}$$

where the second inequality is due to that $U = (U_1, \dots, U_r)$.

By properties of the Sin Θ distance, we have

$$\begin{aligned} \inf_{W \in \mathcal{O}(k,k)} \|U - U^{(i)}W\|_F &\leq \|UU^T - U^{(i)}U^{(i)T}\|_F \\ &\leq 6k^{1.5} \|U_{i\cdot}\|. \end{aligned}$$

□

Lemma 8: Recall the definition of $f_t(\cdot)$ in (10). For any $t > 0$ and any $x, y \in \mathbb{R}^{1 \times k}$, we have $\|f_t(x) - f_t(y)\| \leq \|x - y\|$.

Proof: Without loss of generality, assume $\|x\| \geq \|y\|$. We first state a fact that can be easily verified: if we shrink x until it has the same norm as y , its distance toward y is always decreasing. It implies:

$$\|sx - y\| \leq \|x - y\|, \forall \frac{\|y\|}{\|x\|} \leq s \leq 1. \quad (58)$$

Now we discuss three cases. If $\|x\| \leq t$. Then we have $f_t(x) = x$ and $f_t(y) = y$, and then the equality holds. If $\|y\| \geq t$, we have

$$\begin{aligned} \|f_t(x) - f_t(y)\| &= \left\| \frac{t}{\|x\|}x - \frac{t}{\|y\|}y \right\| = \frac{t}{\|y\|} \left\| \frac{\|y\|}{\|x\|}x - y \right\| \\ &\leq \left\| \frac{\|y\|}{\|x\|}x - y \right\| \leq \|x - y\|, \end{aligned}$$

where the last inequality is due to (58). If $\|x\| \geq t \geq \|y\|$, we have

$$\|f_t(x) - f_t(y)\| = \left\| \frac{t}{\|x\|}x - y \right\| \leq \|x - y\|,$$

where the last inequality is due to (58). □

Lemma 9: Consider any $i \in [n]$. For any $t > 0$, we have

$$\|(I_n - U^*U^{*T})\bar{U}^{(i)}\|_{2,\infty} \leq t + \sqrt{\frac{k}{\beta n}}.$$

Under the assumption $\|\tilde{A}^{(i)} - \mathbb{E}A\| + p \leq (p - q)\beta n/(2k)$, for any $t > \sqrt{k/(\beta n)}$, we have

$$\|(I_n - U^*U^{*T})\bar{U}^{(i)}\|_F \leq \frac{2\sqrt{2}k^{1.5} (\|\tilde{A}^{(i)} - \mathbb{E}A\| + p)}{\beta n(p - q)}.$$

Proof: Note that $\|U^*\|_{2,\infty} \leq \sqrt{k/(\beta n)}$ from Lemma 3 and we have $\lambda_k^* \geq (p - q)\beta n/k$ from Lemma 5. For any $t > 0$, we have $\bar{U}_{j\cdot}^{(i)} = f_t(U_{j\cdot}^{(i)}) = s_j U_{j\cdot}^{(i)}$ for some $s_j \in [0, 1]$ for any $j \in [n]$. Then,

$$\begin{aligned} \|\bar{U}^{(i)}\| &= \max_{x \in \mathbb{R}^k} \|\bar{U}^{(i)}x\| = \max_{x \in \mathbb{R}^k} \sqrt{\sum_{j \in [n]} (\bar{U}_{j\cdot}^{(i)}x)^2} \\ &= \max_{x \in \mathbb{R}^k} \sqrt{\sum_{j \in [n]} s_j^2 (U_{j\cdot}^{(i)}x)^2} \end{aligned}$$

$$\leq \max_{x \in \mathbb{R}^k} \sqrt{\sum_{j \in [n]} (U_{j\cdot}^{(i)}x)^2} = \|U\| = 1.$$

Then,

$$\begin{aligned} \|(I_n - U^*U^{*T})\bar{U}^{(i)}\|_{2,\infty} &\leq \|\bar{U}^{(i)}\|_{2,\infty} + \|U^*U^{*T}\bar{U}^{(i)}\|_{2,\infty} \\ &\leq t + \|U^*\|_{2,\infty} \|U^*T\bar{U}^{(i)}\| \\ &\leq t + \|U^*\|_{2,\infty} \|\bar{U}^{(i)}\| \\ &\leq t + \|U^*\|_{2,\infty} \leq t + \sqrt{\frac{k}{\beta n}}. \end{aligned}$$

On the other hand, we have $\|\tilde{A}^{(i)} - P\| \leq \|\tilde{A}^{(i)} - \mathbb{E}A\| + \|\mathbb{E}A - P\| = \|\tilde{A}^{(i)} - \mathbb{E}A\| + p$. By Davis-Kahan Theorem, when $\|\tilde{A}^{(i)} - \mathbb{E}A\| + p \leq (p - q)\beta n/(2k)$, there exists some orthogonal matrix $W \in \mathcal{O}(k, k)$ such that

$$\|U^* - U^{(i)}W\| \leq \frac{2\|\tilde{A}^{(i)} - P\|}{\lambda_k^*} \leq \frac{2k(\|\tilde{A}^{(i)} - \mathbb{E}A\| + p)}{\beta n(p - q)}.$$

Since $U^* - U^{(i)}W \in \mathbb{R}^{n \times (2k)}$, we have

$$\begin{aligned} \|U^* - U^{(i)}W\|_F &\leq \sqrt{2k} \|U^* - U^{(i)}W\| \\ &\leq \frac{2\sqrt{2}k^{1.5} (\|\tilde{A}^{(i)} - \mathbb{E}A\| + p)}{\beta n(p - q)}. \end{aligned}$$

For any $j \in [n]$ and any $t > \sqrt{k/(\beta n)}$, we have $U_{j\cdot}^* = f_t(U_{j\cdot}^*)$ and

$$\begin{aligned} \|U_{j\cdot}^* - \bar{U}_{j\cdot}^{(i)}W\| &= \|f_t(U_{j\cdot}^*) - f_t(U_{j\cdot}^{(i)}W)\| \\ &\leq \|U_{j\cdot}^* - U_{j\cdot}^{(i)}W\|, \end{aligned}$$

where the last inequality is due to Lemma 8. Hence,

$$\begin{aligned} \|U^* - \bar{U}^{(i)}W\|_F &\leq \|U^* - U^{(i)}W\|_F \\ &\leq \frac{2\sqrt{2}k^{1.5} (\|\tilde{A}^{(i)} - \mathbb{E}A\| + p)}{\beta n(p - q)}. \end{aligned}$$

We get the desired $\|\cdot\|_F$ upper bound with

$$\begin{aligned} \|(I_n - U^*U^{*T})\bar{U}^{(i)}\|_F &= \|(I_n - U^*U^{*T})\bar{U}^{(i)}W\|_F \\ &= \|(I_n - U^*U^{*T})(U^* - \bar{U}^{(i)}W)\|_F \leq \|U^* - \bar{U}^{(i)}W\|_F. \end{aligned}$$

□

Lemma 10: Consider an integer $m > 0$ and independent Bernoulli random variables $\{X_i\}_{i \in [m]}$. Denote $p_{\max} := \max_{i \in [m]} \mathbb{E}X_i$ and assume $p_{\max} \leq 1/2$. For any $s \in \mathbb{R}$ and any $w \in \mathbb{R}^m$, we have

$$\sum_{i \in [m]} \log \mathbb{E} \exp(s w_i (X_i - \mathbb{E}X_i)) \leq p_{\max} s^2 \|w\|^2 e^{|s| \|w\|_\infty}.$$

Proof: Define $f(t; q) := \log(qe^t + (1 - q)) - qt$ to be a function of t where $0 \leq q \leq 1$. Then its first and second derivatives are

$$f'(t; q) = \frac{qe^t}{qe^t + (1 - q)} - q \text{ and } f''(t; q) = \frac{q(1 - q)e^t}{(qe^t + (1 - q))^2}.$$

Note that $f(0; q) = f'(0; q) = 0$ and $f''(t; q) \geq 0$ for all $t \in \mathbb{R}$. If $q \leq 1/2$ is further assumed, we have $f''(t; q) \leq qe^t/(1-q) \leq 2qe^t$ for any t . Then for $t_0 > 0$ and any $t \in [-t_0, t_0]$, we have $f''(t; q) \leq 2qe^{t_0}$ and

$$f(t; q) \leq \frac{1}{2} (2qe^{t_0}) t^2 \leq qe^{t_0} t^2.$$

Hence,

$$\begin{aligned} \sum_{i \in [m]} \log \mathbb{E} \exp(s w_i (X_i - \mathbb{E} X_i)) &= \sum_{i \in [m]} f(s w_i; \mathbb{E} X_i) \\ &\leq \sum_{i \in [m]} (\mathbb{E} X_i) e^{|s| \|w\|_\infty} s^2 w_i^2 \leq p_{\max} s^2 \|w\|^2 e^{|s| \|w\|_\infty}. \end{aligned}$$

□

Lemma 11: Consider any $s > 0$ and any integer $m > 0$. For any $\{a_i\}_{i \in [m]}$ such that $a_i \geq 0, \forall i \in [m]$, we have

$$\left(\sum_{i \in [m]} a_i \right)^2 \mathbb{I} \left\{ \sum_{i \in [m]} a_i \geq s \right\} \leq m^2 \sum_{i \in [m]} a_i^2 \mathbb{I} \left\{ a_i \geq \frac{s}{m} \right\}.$$

Proof: We have

$$\begin{aligned} &\left(\sum_{i \in [m]} a_i \right)^2 \mathbb{I} \left\{ \sum_{i \in [m]} a_i \geq s \right\} \\ &\leq \left(\sum_{i \in [m]} a_i \right)^2 \sum_{l \in [m]} \mathbb{I} \left\{ \sum_{i \in [m]} a_i \geq s, a_l \geq \max_{j: j \neq l} a_j \right\} \\ &= \sum_{l \in [m]} \left(a_l + \sum_{j: j \neq l} a_j \right)^2 \mathbb{I} \left\{ \sum_{i \in [m]} a_i \geq s, a_l \geq \max_{j: j \neq l} a_j \right\} \\ &\leq \sum_{l \in [m]} (m a_l)^2 \mathbb{I} \left\{ \sum_{i \in [m]} a_i \geq s, a_l \geq \max_{j: j \neq l} a_j \right\} \\ &\leq \sum_{l \in [m]} (m a_l)^2 \mathbb{I} \left\{ a_l \geq \frac{s}{m} \right\}. \end{aligned}$$

□

Lemma 12: Consider any two scalars $s, t > 8$. Then

$$\sum_{j=1}^{\infty} (j+1)^2 \exp \left(-\frac{\frac{1}{2} j^2}{\frac{1}{s} + \frac{1}{3t}} \right) \leq 8 \exp \left(-\frac{\frac{1}{2}}{\frac{1}{s} + \frac{1}{3t}} \right).$$

Proof: For any $j \geq 1$, we have

$$\begin{aligned} \frac{\frac{1}{2}(j+1)^2}{\frac{1}{s} + \frac{1}{3} \frac{j+1}{t}} &= \frac{(j+1)^2}{j^2} \frac{\frac{1}{s} + \frac{1}{3} \frac{j}{t}}{\frac{1}{s} + \frac{1}{3} \frac{j+1}{t}} = \frac{(j+1)^2}{j^2} \frac{1}{1 + \frac{\frac{1}{3t}}{\frac{1}{s} + \frac{1}{3t}}} \\ &= \frac{(j+1)^2}{j^2} \frac{1}{1 + \frac{1}{\frac{3t}{s} + j}} \geq \frac{(j+1)^2}{j^2} \frac{1}{1 + \frac{1}{j}} \geq \frac{j+1}{j}. \end{aligned}$$

Hence,

$$\frac{\frac{1}{2} j^2}{\frac{1}{s} + \frac{1}{3t}} \geq \frac{j}{j-1} \frac{j-1}{j-2} \cdots \frac{2}{1} \frac{1}{\frac{1}{s} + \frac{1}{3t}} = j \frac{1}{\frac{1}{s} + \frac{1}{3t}}.$$

As a result,

$$\sum_{j=1}^{\infty} (j+1)^2 \exp \left(-\frac{\frac{1}{2} j^2}{\frac{1}{s} + \frac{1}{3t}} \right)$$

$$\leq \sum_{j=1}^{\infty} (j+1)^2 \exp \left(-j \frac{\frac{1}{2}}{\frac{1}{s} + \frac{1}{3t}} \right).$$

Since $s, t > 8$, we have $\frac{\frac{1}{2}}{\frac{1}{s} + \frac{1}{3t}} > 2$. Then the first term in $\left\{ \exp \left(-j \frac{\frac{1}{2}}{\frac{1}{s} + \frac{1}{3t}} \right) \right\}_{j \geq 1}$ dominates. As a result, the above display is upper bounded by $8 \exp \left(-\frac{\frac{1}{2}}{\frac{1}{s} + \frac{1}{3t}} \right)$. □

Lemma 13: Recall the definition of $J_{m_1, m_2, p, q}$ in (7). For any positive integers m_1, m_2 and any p, q such that $0 < q < p \leq 1/2$, we have

$$m_2 \frac{(p-q)^2}{8p} \leq J_{m_1, m_2, p, q} \leq (m_1 + m_2) \frac{(p-q)^2}{4q},$$

and

$$J_{m_1, m_2+1, p, q} \leq J_{m_1, m_2, p, q} + \frac{(p-q)^2}{4q}.$$

In addition, define

$$\begin{aligned} t^* := \operatorname{argmax}_t &\left((m_1 - m_2) t \frac{p+q}{2} - m_1 \log(qe^t + 1 - q) \right. \\ &\left. - m_2 \log(pe^{-t} + 1 - p) \right). \end{aligned}$$

We have

$$0 < t^* \leq \frac{p-q}{q}. \quad (59)$$

If $p \leq 1/10$ is further assumed, we have

$$J_{m_1, m_2, p, q} \wedge J_{m_2, m_1, p, q} \leq (m_2 \vee m_1) \frac{4(p-q)^2}{3p}.$$

Proof: We introduce auxiliary functions:

$$\begin{aligned} g_1(t) &:= t \frac{p+q}{2} - \log(qe^t + 1 - q) \\ g_2(t) &:= -t \frac{p+q}{2} - \log(pe^{-t} + 1 - p) \\ f(t) &:= m_1 g_1(t) + m_2 g_2(t). \end{aligned}$$

Then $J_{m_1, m_2, p, q} = \max_t f(t) = f(t^*)$. Through calculation, we have

$$\begin{aligned} g'_1(t) &= - \left(1 - \frac{p+q}{2} - \frac{1-q}{qe^t + 1 - q} \right) \\ g'_2(t) &= 1 - \frac{p+q}{2} - \frac{1-p}{pe^{-t} + 1 - p} \\ g''_1(t) &= - \frac{(1-q)qe^t}{(qe^t + 1 - q)^2} \\ g''_2(t) &= - \frac{(1-p)pe^{-t}}{(pe^{-t} + 1 - p)^2} \\ f'(t) &= m_1 g'_1(t) + m_2 g'_2(t) \\ f''(t) &= m_1 g''_1(t) + m_2 g''_2(t). \end{aligned}$$

Note that $g''_1(t), g''_2(t), f''(t) < 0$ for all $t \in \mathbb{R}$ which implies $g'_1(t), g'_2(t), f'(t)$ are all decreasing functions. Define

$$t_1 := \log \left(\frac{1-q}{1 - \frac{p+q}{2}} \frac{\frac{p+q}{2}}{q} \right) = \log \left(1 + \frac{\frac{p-q}{2}}{(1 - \frac{p+q}{2})q} \right)$$

and

$$t_2 := \log \left(\frac{1 - \frac{p+q}{2}}{1-p} \frac{p}{\frac{p+q}{2}} \right) = \log \left(1 + \frac{\frac{p-q}{2}}{(1-p)\frac{p+q}{2}} \right),$$

such that $g'_1(t_1) = 0$ and $g'_2(t_2) = 0$. Under the assumption that $0 < q < p \leq 1/2$, we have $(1 - \frac{p+q}{2})q \leq (1-p)\frac{p+q}{2}$ and consequently $0 < t_2 \leq t_1$. Using the fact $f'(t) = m_1g'_1(t) + m_2g'_2(t)$ and that g'_1, g'_2 are decreasing function, we have $f'(t_2) = m_1g'_1(t_2) \geq m_1g'_1(t_1) = 0$ and $f'(t_1) = m_2g'_2(t_1) \leq m_2g'_2(t_2) = 0$. That $f'(t)$ is a decreasing function leads to

$$0 < t_2 \leq t^* \leq t_1. \quad (60)$$

Let us first study $g_2(t_2)$ and $g_1(t_1)$ which are important quantities for further analysis. Through calculation, we can show that $g_2(t_2)$ can be simplified into

$$g_2(t_2) = \left(1 - \frac{p+q}{2} \right) \log \frac{1 - \frac{p+q}{2}}{1-p} + \frac{p+q}{2} \log \frac{\frac{p+q}{2}}{p}.$$

Define $h(\delta) := (p+\delta) \log \frac{p+\delta}{p} + (1-(p+\delta)) \log \frac{1-(p+\delta)}{1-p}$. We have $h(0) = 0$ and $h''(\delta) = \frac{1}{(p+\delta)(1-(p+\delta))}$. Since $0 < q < p \leq 1/2$, we have $h''(\delta) \geq \frac{1}{p(1-p)}$ for any $-\frac{p-q}{2} \leq \delta \leq 0$ and consequently,

$$g_2(t_2) = h\left(-\frac{p-q}{2}\right) \geq \frac{1}{2} \frac{1}{p(1-p)} \left(\frac{p-q}{2}\right)^2 \geq \frac{(p-q)^2}{8p}. \quad (61)$$

Similarly, we have

$$\begin{aligned} g_2(t_2) &\leq \frac{1}{2} \left(\max_{-\frac{p-q}{2} \leq \delta \leq 0} h''(\delta) \right) \left(\frac{p-q}{2}\right)^2 \\ &\leq \frac{1}{2} \frac{1}{\frac{p+q}{2}(1-\frac{p+q}{2})} \left(\frac{p-q}{2}\right)^2 \leq \frac{(p-q)^2}{4q}. \end{aligned} \quad (62)$$

Using the same argument, we have

$$\begin{aligned} g_1(t_1) &= \left(1 - \frac{p+q}{2} \right) \log \frac{1 - \frac{p+q}{2}}{1-q} + \frac{p+q}{2} \log \frac{\frac{p+q}{2}}{q} \\ &\leq \frac{1}{2} \frac{1}{q(1-q)} \left(\frac{p-q}{2}\right)^2 \leq \frac{(p-q)^2}{4q}. \end{aligned} \quad (63)$$

Now we are ready to establish the bounds for $J_{m_1, m_2, p, q}$. For the upper bound, we have

$$J_{m_1, m_2, p, q} \leq m_1 g_1(t_1) + m_2 g_2(t_2) \leq (m_1 + m_2) \frac{(p-q)^2}{4q},$$

by (62) and (63). For the lower bound, note that $g_1(t_2) \geq g_1(0) = 0$. Then by (61), we have

$$J_{m_1, m_2, p, q} \geq f(t_2) \geq m_2 g_2(t_2) \geq m_2 \frac{(p-q)^2}{8p}.$$

From (60) and the definition of t_1 , we also have

$$0 < t^* \leq \frac{\frac{p-q}{2}}{(1-\frac{p+q}{2})q} \leq \frac{p-q}{q}, \quad (64)$$

where the last inequality is due to that $0 < p, q \leq 1/2$. Note that t^* is a function of m_1, m_2 and (64) still holds if we vary

values of m_1, m_2 . This implies the maximizer of $m_1 g_1(t) + (m_2 + 1)g_2(t)$ is also within $[0, \frac{p-q}{q}]$. Hence,

$$\begin{aligned} J_{m_1, m_2+1, p, q} &= \max_{t \in [0, \frac{p-q}{q}]} (m_1 g_1(t) + (m_2 + 1)g_2(t)) \\ &= \max_{t \in [0, \frac{p-q}{q}]} (f(t) + g_2(t)) \\ &\leq \max_{t \in [0, \frac{p-q}{q}]} (f(t) + g_2(t_2)) \\ &= J_{m_1, m_2, p, q} + g_2(t_2) \\ &\leq J_{m_1, m_2, p, q} + \frac{(p-q)^2}{4q}, \end{aligned}$$

where the last inequality is due to (62).

In the last part of the proof, we are going to derive an improved upper bound for $J_{m_1, m_2, p, q} \wedge J_{m_2, m_1, p, q}$ under an additional assumption that $p \leq 1/10$. From (60), we have $g_1(t^*) \geq g_1(0) = 0$. Let us first consider the case that $m_1 \leq m_2$. Then

$$\begin{aligned} J_{m_1, m_2, p, q} &= f(t^*) = m_1 g_1(t^*) + m_2 g_2(t^*) \\ &\leq m_2 g_1(t^*) + m_2 g_2(t^*) \\ &\leq m_2 \max_{t \geq 0} (g_1(t) + g_2(t)) \leq m_2 \frac{4(p-q)^2}{3p}, \end{aligned}$$

where the last equation is due to Lemma 14. If $m_1 > m_2$ instead, by the same argument, we have $J_{m_2, m_1, p, q} \leq m_1 \frac{4(p-q)^2}{3p}$. Hence, $J_{m_1, m_2, p, q} \wedge J_{m_2, m_1, p, q} \leq (m_2 \vee m_1) \frac{4(p-q)^2}{3p}$ holds for both cases. \square

Lemma 14: Consider any $0 < q < p < 1$. Define $I_{p, q} := -2 \log(\sqrt{pq} + \sqrt{(1-p)(1-q)})$. Then $I_{p, q} > 0$ and

$$I_{p, q} = \max_t (-\log(pe^{-t} + 1 - p) - \log(qe^t + 1 - q)).$$

If $p \leq 1/10$ is further assumed, we have $I_{p, q} \leq \frac{4(p-q)^2}{3p}$.

Proof: The equation for $I_{p, q}$ is by direct calculation and its proof is omitted here. For the upper bound of $I_{p, q}$, we have

$$\begin{aligned} I_{p, q} &= -\log \left(1 - (p+q) + 2\sqrt{pq} \left(\sqrt{pq} + \sqrt{(1-p)(1-q)} \right) \right) \\ &= -\log \left(1 - (\sqrt{p} - \sqrt{q})^2 \right. \\ &\quad \left. - 2\sqrt{pq} \left(1 - \sqrt{pq} - \sqrt{(1-p)(1-q)} \right) \right) \\ &= -\log \left(1 - (\sqrt{p} - \sqrt{q})^2 \right. \\ &\quad \left. - 2\sqrt{pq} \frac{(\sqrt{p} - \sqrt{q})^2}{1 - \sqrt{pq} + \sqrt{(1-p)(1-q)}} \right) \\ &= -\log \left(1 - \left(1 + \frac{2\sqrt{pq}}{1 - \sqrt{pq} + \sqrt{(1-p)(1-q)}} \right) \right. \\ &\quad \left. \times (\sqrt{p} - \sqrt{q})^2 \right). \end{aligned}$$

Note that $-\log(1-x) \leq 1.2x, \forall 0 \leq x \leq 0.3$. When $p \leq 1/2$,

$$\left(1 + \frac{2\sqrt{pq}}{1 - \sqrt{pq} + \sqrt{(1-p)(1-q)}} \right) (\sqrt{p} - \sqrt{q})^2$$

$$\leq 2(\sqrt{p} - \sqrt{q})^2 \leq 2p.$$

Hence, if we further assume $p \leq 1/10$, the above display is smaller than 0.3, and we have

$$\begin{aligned} I_{p,q} &\leq 1.2 \left(1 + \frac{2\sqrt{pq}}{1 - \sqrt{pq} + \sqrt{(1-p)(1-q)}} \right) (\sqrt{p} - \sqrt{q})^2 \\ &\leq 1.2 \left(1 + \frac{2p}{1 - 0.1 + 0.9} \right) \frac{(p-q)^2}{p} \\ &\leq \frac{4(p-q)^2}{3p}. \end{aligned}$$

□

Lemma 15: Consider any two integers $m_1, m_2 > 0$ and any p, q such that $0 < q < p \leq 1/2$. Let $\{X_i\}_{i \in [m_1]} \stackrel{iid}{\sim} \text{Bernoulli}(q)$, $\{Y_j\}_{j \in [m_2]} \stackrel{iid}{\sim} \text{Bernoulli}(p)$ and assume they are independent of each other. Consider any ρ such that $0 \leq \rho \leq 1$. We have

$$\begin{aligned} &\mathbb{P} \left(\sum_{i \in [m_1]} (X_i - q) - \sum_{j \in [m_2]} (Y_j - p) \right. \\ &\quad \left. \geq \frac{1-\rho}{2} (m_1 + m_2) (p - q) \right) \\ &\leq \exp \left(-J_{m_1, m_2, p, q} + \frac{\rho}{2} (m_1 + m_2) \frac{(p-q)^2}{q} \right). \end{aligned}$$

and

$$\begin{aligned} &\mathbb{P} \left(\sum_{i \in [m_1]} (X_i - q) - \sum_{j \in [m_2]} (Y_j - p) \right. \\ &\quad \left. \geq \frac{1+\rho}{2} (m_1 + m_2) (p - q) \right) \\ &\geq \exp \left(-J_{m_1, m_2, p, q} - \rho (m_1 + m_2) \frac{(p-q)^2}{q} \right. \\ &\quad \left. - 4\sqrt{\frac{(m_1 + m_2)(p-q)^2 p}{q^2}} \right) \left(\frac{1}{4} - \sqrt{\frac{2}{(m_1 + m_2)q}} \right). \end{aligned}$$

Proof: We first prove the upper bound. By Chernoff bound, we have

$$\begin{aligned} &\log \mathbb{P} \left(\sum_{i \in [m_1]} (X_i - q) - \sum_{j \in [m_2]} (Y_j - p) \right. \\ &\quad \left. \geq \frac{1-\rho}{2} (m_1 + m_2) (p - q) \right) \\ &= \log \mathbb{P} \left(\sum_{i \in [m_1]} X_i - \sum_{j \in [m_2]} Y_j \geq (m_1 - m_2) \frac{p+q}{2} \right. \\ &\quad \left. - \frac{\rho}{2} (m_1 + m_2) (p - q) \right) \\ &\leq \min_{t>0} \left(- (m_1 - m_2) t \frac{p+q}{2} + \frac{\rho}{2} (m_1 + m_2) (p - q) t \right. \\ &\quad \left. + m_1 \log (qe^t + 1 - q) + m_2 \log (pe^{-t} + 1 - p) \right) \end{aligned}$$

$$\begin{aligned} &= - \max_{t>0} \left((m_1 - m_2) t \frac{p+q}{2} - \frac{\rho}{2} (m_1 + m_2) (p - q) t \right. \\ &\quad \left. - m_1 \log (qe^t + 1 - q) - m_2 \log (pe^{-t} + 1 - p) \right) \\ &\leq - \left(J_{m_1, m_2, p, q} - \frac{\rho}{2} (m_1 + m_2) (p - q) t^* \right) \\ &\leq -J_{m_1, m_2, p, q} + \frac{\rho}{2} (m_1 + m_2) \frac{(p-q)^2}{q}. \end{aligned}$$

where t^* is defined in the statement of Lemma 13. Here the second to last inequality holds as $t^* > 0$ according to (59) and the last inequality holds as $t^* \leq \frac{p-q}{q}$ according to (59) as well.

We next prove the lower bound. Define

$$\begin{aligned} \tilde{f}(t) &:= (m_1 - m_2) t \frac{p+q}{2} + \frac{\rho}{2} (m_1 + m_2) (p - q) t \\ &\quad - m_1 \log (qe^t + 1 - q) - m_2 \log (pe^{-t} + 1 - p) \\ &= m_1 \left(t \left(\frac{p+q}{2} + \frac{\rho(p-q)}{2} \right) - \log (qe^t + 1 - q) \right) \\ &\quad + m_2 \left(- \left(\frac{p+q}{2} - \frac{\rho(p-q)}{2} \right) - \log (pe^{-t} + 1 - p) \right), \end{aligned}$$

$\tilde{t} := \arg\max_t \tilde{f}(t)$, and $\tilde{J}_{m_1, m_2, p, q, \rho} := \tilde{f}(\tilde{t}) = \max_t \tilde{f}(t)$. Define

$$\begin{aligned} \tilde{t}_1 &:= \log \left(\frac{1-q}{1 - \left(\frac{p+q}{2} + \frac{\rho(p-q)}{2} \right)} \frac{\frac{p+q}{2} + \frac{\rho(p-q)}{2}}{q} \right) \\ &= \log \left(1 + \frac{(1+\rho) \frac{p-q}{2}}{\left(1 - \left(\frac{p+q}{2} + \frac{\rho(p-q)}{2} \right) \right) q} \right) \end{aligned}$$

and

$$\begin{aligned} \tilde{t}_2 &:= \log \left(\frac{1 - \left(\frac{p+q}{2} - \frac{\rho(p-q)}{2} \right)}{1-p} \frac{p}{\frac{p+q}{2} - \frac{\rho(p-q)}{2}} \right) \\ &= \log \left(1 + \frac{(1+\rho) \frac{p-q}{2}}{(1-p) \left(\frac{p+q}{2} - \frac{\rho(p-q)}{2} \right)} \right). \end{aligned}$$

Following the same argument used to derive (60) in the proof of Lemma 13, we have \tilde{t} sandwiched between \tilde{t}_1 and \tilde{t}_2 . Hence, similar to (64), we have

$$0 < \tilde{t} \leq \tilde{t}_1 \vee \tilde{t}_2 \leq \frac{2(p-q)}{q}, \quad (65)$$

where we use the assumption $0 \leq \rho \leq 1$ and $0 < q < p \leq 1/2$. In addition, we have

$$\begin{aligned} 1 &\leq e^{\tilde{t}} \leq e^{\tilde{t}_1} \vee e^{\tilde{t}_2} \\ &\leq 1 + \frac{(1+\rho) \frac{p-q}{2}}{\left(1 - \left(\frac{p+q}{2} + \frac{\rho(p-q)}{2} \right) \right) q} \vee \frac{(1+\rho) \frac{p-q}{2}}{(1-p) \left(\frac{p+q}{2} - \frac{\rho(p-q)}{2} \right)} \\ &\leq 1 + \frac{2(p-q)}{q} \leq \frac{2p}{q}. \end{aligned} \quad (66)$$

Using (65), we have

$$\tilde{J}_{m_1, m_2, p, q, \rho}$$

$$\begin{aligned}
&= (m_1 - m_2) \tilde{t} \frac{p+q}{2} + \frac{\rho}{2} (m_1 + m_2) (p - q) \tilde{t} \\
&\quad - m_1 \log(qe^{\tilde{t}} + 1 - q) - m_2 \log(pe^{-\tilde{t}} + 1 - p) \\
&\leq \max_t \left((m_1 - m_2) t \frac{p+q}{2} - m_1 \log(qe^t + 1 - q) \right. \\
&\quad \left. - m_2 \log(pe^{-t} + 1 - p) \right) + \frac{\rho}{2} (m_1 + m_2) (p - q) \tilde{t} \\
&= J_{m_1, m_2, p, q} + \frac{\rho}{2} (m_1 + m_2) (p - q) \tilde{t} \\
&\leq J_{m_1, m_2, p, q} + \rho (m_1 + m_2) \frac{(p - q)^2}{q}. \tag{67}
\end{aligned}$$

Define $M := \tilde{t} \left(\frac{p+q}{2} + \frac{\rho(p-q)}{2} \right) - \log(qe^{\tilde{t}} + 1 - q)$ and $N := -\tilde{t} \left(\frac{p+q}{2} - \frac{\rho(p-q)}{2} \right) - \log(pe^{-\tilde{t}} + 1 - p)$. We are going to use the Cramér-Chernoff argument to establish the lower bound. We have

$$\begin{aligned}
&\mathbb{P} \left(\sum_{i \in [m_1]} (X_i - q) - \sum_{j \in [m_2]} (Y_j - p) \right. \\
&\quad \left. \geq \frac{1+\rho}{2} (m_1 + m_2) (p - q) \right) \\
&= \mathbb{P} \left(\sum_{i \in [m_1]} \left(X_i - \left(\frac{p+q}{2} + \frac{\rho(p-q)}{2} \right) \right) \right. \\
&\quad \left. - \sum_{j \in [m_2]} \left(Y_j - \left(\frac{p+q}{2} - \frac{\rho(p-q)}{2} \right) \right) \geq 0 \right) \\
&\geq \mathbb{P} \left((m_1 + m_2) \delta \geq \sum_{i \in [m_1]} \left(X_i - \left(\frac{p+q}{2} + \frac{\rho(p-q)}{2} \right) \right) \right. \\
&\quad \left. - \sum_{j \in [m_2]} \left(Y_j - \left(\frac{p+q}{2} - \frac{\rho(p-q)}{2} \right) \right) \geq 0 \right) \\
&= \sum_{(x,y) \in \mathcal{X}} \left(\prod_{i \in [m_1]} h_1(x_i) \right) \left(\prod_{j \in [m_2]} h_2(y_j) \right).
\end{aligned}$$

In the above display, $\delta > 0$ is some quantity whose value will be given later. The set $\mathcal{X} := \{(x, y) : x \in \mathbb{R}^{m_1}, y \in \mathbb{R}^{m_2}, (m_1 + m_2) \delta \geq \sum_{i \in [m_1]} x_i - \sum_{j \in [m_2]} y_j \geq 0\}$ and $h_1(\cdot), h_2(\cdot)$ are defined to be the probability mass functions of $X_1 - \left(\frac{p+q}{2} + \frac{\rho(p-q)}{2} \right)$ and $Y_1 - \left(\frac{p+q}{2} - \frac{\rho(p-q)}{2} \right)$, respectively. We further define

$$\begin{aligned}
M &:= \mathbb{E} \exp \left(\tilde{t} \left(X_1 - \left(\frac{p+q}{2} + \frac{\rho(p-q)}{2} \right) \right) \right) \\
&= qe^{\tilde{t}(1 - (\frac{p+q}{2} + \frac{\rho(p-q)}{2}))} + (1 - q)e^{-\tilde{t}(\frac{p+q}{2} + \frac{\rho(p-q)}{2})}, \\
N &:= \mathbb{E} \exp \left(-\tilde{t} \left(Y_1 - \left(\frac{p+q}{2} - \frac{\rho(p-q)}{2} \right) \right) \right) \\
&= pe^{-\tilde{t}(1 - (\frac{p+q}{2} - \frac{\rho(p-q)}{2}))} + (1 - p)e^{\tilde{t}(\frac{p+q}{2} - \frac{\rho(p-q)}{2})}.
\end{aligned}$$

Then

$$\mathbb{P} \left(\sum_{i \in [m_1]} (X_i - q) - \sum_{j \in [m_2]} (Y_j - p) \right.$$

$$\begin{aligned}
&\geq \frac{1+\rho}{2} (m_1 + m_2) (p - q) \Big) \\
&= M^{m_1} N^{m_2} \sum_{(x,y) \in \mathcal{X}} \left(\prod_{i \in [m_1]} \frac{\exp(\tilde{t} x_i) h_1(x_i)}{\exp(\tilde{t} x_i) M} \right) \\
&\quad \times \left(\prod_{j \in [m_2]} \frac{\exp(-\tilde{t} y_j) h_2(y_j)}{\exp(-\tilde{t} y_j) N} \right) \\
&\geq \frac{M^{m_1} N^{m_2}}{\exp((m_1 + m_2) \tilde{t} \delta)} \sum_{(x,y) \in \mathcal{X}} \left(\prod_{i \in [m_1]} \frac{\exp(\tilde{t} x_i) h_1(x_i)}{M} \right) \\
&\quad \times \left(\prod_{j \in [m_2]} \frac{\exp(-\tilde{t} y_j) h_2(y_j)}{N} \right),
\end{aligned}$$

where the last inequality is due to that $\sum_{i \in [m_1]} x_i - \sum_{j \in [m_2]} y_j \leq (m_1 + m_2) \delta$ as $(x, y) \in \mathcal{X}$. Define $h'_1(w) := \frac{\exp(\tilde{t} w) h_1(w)}{M}$ and $h'_2(w) := \frac{\exp(-\tilde{t} w) h_2(w)}{N}$. Since $\sum_w h'_1(w) = \sum_w h'_2(w) = 1$, they are both probability mass functions. Let U_1, \dots, U_{m_1} be i.i.d. random variables distributed according to $h'_1(\cdot)$ and let V_1, \dots, V_{m_2} be i.i.d. random variables distributed according to $h'_2(\cdot)$. We further assume they are independent of each other. Then

$$\begin{aligned}
&\mathbb{P} \left(\sum_{i \in [m_1]} (X_i - q) - \sum_{j \in [m_2]} (Y_j - p) \right. \\
&\quad \left. \geq \frac{1+\rho}{2} (m_1 + m_2) (p - q) \right) \\
&\geq \frac{M^{m_1} N^{m_2}}{\exp((m_1 + m_2) \tilde{t} \delta)} \\
&\quad \times \sum_{(x,y) \in \mathcal{X}} \left(\prod_{i \in [m_1]} h'_1(x_i) \right) \left(\prod_{j \in [m_2]} h'_2(y_j) \right) \\
&= \frac{M^{m_1} N^{m_2}}{\exp((m_1 + m_2) \tilde{t} \delta)} \\
&\quad \times \mathbb{P} \left((m_1 + m_2) \delta \geq \sum_{i \in [m_1]} U_i - \sum_{j \in [m_2]} V_j \geq 0 \right).
\end{aligned}$$

Note that $M^{m_1} N^{m_2} = \exp(-\tilde{J}_{m_1, m_2, p, q, \rho}) \geq \exp(-J_{m_1, m_2, p, q} - \rho(m_1 + m_2) \frac{(p-q)^2}{q})$ where we use (67). Hence,

$$\begin{aligned}
&\mathbb{P} \left(\sum_{i \in [m_1]} (X_i - q) - \sum_{j \in [m_2]} (Y_j - p) \right. \\
&\quad \left. \geq \frac{1+\rho}{2} (m_1 + m_2) (p - q) \right) \\
&\geq \exp \left(-J_{m_1, m_2, p, q} - \rho(m_1 + m_2) \frac{(p-q)^2}{q} \right. \\
&\quad \left. - (m_1 + m_2) \tilde{t} \delta \right) \mathbb{P} \left((m_1 + m_2) \delta \right.
\end{aligned}$$

$$\geq \left(\sum_{i \in [m_1]} U_i - \sum_{j \in [m_2]} V_j \right) \geq 0 \Bigg).$$

Note that

$$\mathbb{E} \left(\sum_{i \in [m_1]} U_i - \sum_{j \in [m_2]} V_j \right) = m_1 \mathbb{E} U_1 - m_2 \mathbb{E} V_1 = \tilde{f}'(\tilde{t}) = 0,$$

where the second equation is by calculation and the third equation is by the fact that \tilde{t} is the maximizer of \tilde{f} . Then $\sum_{i \in [m_1]} U_i - \sum_{j \in [m_2]} V_j = \sum_{i \in [m_1]} (U_i - \mathbb{E} U_i) - \sum_{j \in [m_2]} (V_j - \mathbb{E} V_j)$. Through calculation, we have

$$\begin{aligned} \mathbb{E} |U_1 - \mathbb{E} U_1|^2 &= \text{Var}(U_1) = \frac{(1-q)qe^{\tilde{t}}}{(qe^{\tilde{t}} + 1 - q)^2}, \\ \mathbb{E} |V_1 - \mathbb{E} V_1|^2 &= \text{Var}(V_1) = \frac{(1-p)pe^{-\tilde{t}}}{(pe^{-\tilde{t}} + 1 - p)^2}. \end{aligned}$$

In addition, we have $|U_1 - \mathbb{E} U_1| \leq 1$ and consequently $\mathbb{E} |U_1 - \mathbb{E} U_1|^3 \leq \mathbb{E} |U_1 - \mathbb{E} U_1|^2$. Similarly, we have $\mathbb{E} |V_1 - \mathbb{E} V_1|^3 \leq \mathbb{E} |V_1 - \mathbb{E} V_1|^2$. Then by Berry-Essen Theorem, we have

$$\begin{aligned} \mathbb{P} \left((m_1 + m_2) \delta \geq \left(\sum_{i \in [m_1]} U_i - \sum_{j \in [m_2]} V_j \right) \geq 0 \right) \\ &= \mathbb{P} \left(\frac{(m_1 + m_2) \delta}{\sqrt{m_1 \text{Var}(U_1) + m_2 \text{Var}(U_2)}} \geq \frac{\sum_{i \in [m_1]} U_i - \sum_{j \in [m_2]} V_j}{\sqrt{m_1 \text{Var}(U_1) + m_2 \text{Var}(U_2)}} \geq 0 \right) \\ &\geq \exp \left(\frac{(m_1 + m_2) \delta}{\sqrt{m_1 \text{Var}(U_1) + m_2 \text{Var}(U_2)}} \geq \mathcal{N}(0, 1) \geq 0 \right) \\ &\quad - \frac{1}{\sqrt{m_1 \text{Var}(U_1) + m_2 \text{Var}(U_2)}}. \end{aligned}$$

Through direct calculation, we have

$$\begin{aligned} &m_1 \text{Var}(U_1) + m_2 \text{Var}(U_2) \\ &= m_1 \frac{(1-q)qe^{\tilde{t}}}{(qe^{\tilde{t}} + 1 - q)^2} + m_2 \frac{(1-p)pe^{-\tilde{t}}}{(pe^{-\tilde{t}} + 1 - p)^2} \\ &\leq m_1 \frac{qe^{\tilde{t}}}{1-q} + m_2 \frac{pe^{-\tilde{t}}}{1-p} \\ &\leq m_1 \frac{q}{1-q} \frac{2p}{q} + m_2 \frac{p}{1-p} \\ &\leq 4(m_1 + m_2)p, \end{aligned}$$

where the second inequality is due to (66). Note that as a function of t , $\frac{(1-q)qe^t}{(qe^t + 1 - q)^2}$ first increase and then decrease when $t \geq 0$ and grows. By (66) again, we have

$$\frac{(1-q)qe^{\tilde{t}}}{(qe^{\tilde{t}} + 1 - q)^2} \geq \frac{(1-q)q}{(q + 1 - q)^2} \wedge \frac{(1-q)q \frac{2p}{q}}{\left(q \frac{2p}{q} + 1 - q\right)^2} \geq \frac{q}{2}.$$

The same lower bound holds for $\frac{(1-p)pe^{-\tilde{t}}}{(pe^{-\tilde{t}} + 1 - p)^2}$. As a result,

$$m_1 \text{Var}(U_1) + m_2 \text{Var}(U_2) \geq (m_1 + m_2) \frac{q}{2}.$$

Taking $\delta = 2\sqrt{\frac{p}{m_1 + m_2}}$, we have

$$\begin{aligned} \mathbb{P} \left((m_1 + m_2) \delta \geq \left(\sum_{i \in [m_1]} U_i - \sum_{j \in [m_2]} V_j \right) \geq 0 \right) \\ &\geq \mathbb{P} (1 \geq \mathcal{N}(0, 1) \geq 0) - \sqrt{\frac{2}{(m_1 + m_2)q}} \\ &\geq \frac{1}{4} - \sqrt{\frac{2}{(m_1 + m_2)q}}. \end{aligned}$$

With this choice of δ , using (65), we have

$$\begin{aligned} \mathbb{P} \left(\sum_{i \in [m_1]} (X_i - q) - \sum_{j \in [m_2]} (Y_j - p) \geq \frac{1+\rho}{2} (m_1 + m_2) (p - q) \right) \\ &\geq \exp \left(-J_{m_1, m_2, p, q} - \rho (m_1 + m_2) \frac{(p - q)^2}{q} \right. \\ &\quad \left. - 4\sqrt{\frac{(m_1 + m_2)(p - q)^2 p}{q^2}} \right) \left(\frac{1}{4} - \sqrt{\frac{2}{(m_1 + m_2)q}} \right). \end{aligned}$$

□

REFERENCES

- [1] E. Abbe, "Community detection and stochastic block models: Recent developments," *J. Mach. Learn. Res.*, vol. 18, no. 1, pp. 6446–6531, Jan. 2017.
- [2] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proc. Nat. Acad. Sci. USA*, vol. 99, no. 12, pp. 7821–7826, Jun. 2002.
- [3] Y. Zhao, E. Levina, and J. Zhu, "Community extraction for social networks," *Proc. Nat. Acad. Sci. USA*, vol. 108, no. 18, pp. 7321–7326, May 2011.
- [4] K. Rohe, S. Chatterjee, and B. Yu, "Spectral clustering and the high-dimensional stochastic blockmodel," *Ann. Statist.*, vol. 39, no. 4, pp. 1878–1915, 2011.
- [5] U. von Luxburg, M. Belkin, and O. Bousquet, "Consistency of spectral clustering," *Ann. Statist.*, vol. 36, pp. 555–586, Apr. 2008.
- [6] J. Jin, Z. T. Ke, and S. Luo, "Mixed membership estimation for social networks," *J. Econometrics*, vol. 239, no. 2, Feb. 2024, Art. no. 105369.
- [7] J. Jin, "Fast community detection by SCORE," *Ann. Statist.*, vol. 43, no. 1, pp. 57–89, Feb. 2015.
- [8] L. Lei, X. Li, and X. Lou, "Consistency of spectral clustering on hierarchical stochastic block models," 2020, *arXiv:2004.14531*.
- [9] S. Paul and Y. Chen, "Spectral and matrix factorization methods for consistent community detection in multi-layer networks," *Ann. Statist.*, vol. 48, no. 1, pp. 230–250, 2020.
- [10] L. Gulikers, M. Lelarge, and L. Massoulié, "A spectral method for community detection in moderately sparse degree-corrected stochastic block models," *Adv. Appl. Probab.*, vol. 49, no. 3, pp. 686–721, Sep. 2017.
- [11] Y. Zhang, E. Levina, and J. Zhu, "Detecting overlapping communities in networks using spectral methods," *SIAM J. Math. Data Sci.*, vol. 2, no. 2, pp. 265–283, Jan. 2020.
- [12] D. Ghoshdastidar and A. Dukkipati, "Consistency of spectral hypergraph partitioning under planted partition model," *Ann. Statist.*, vol. 45, no. 1, pp. 289–315, Feb. 2017.
- [13] Z. T. Ke and J. Wang, "Optimal network membership estimation under severe degree heterogeneity," 2022, *arXiv:2204.12087*.
- [14] Z. T. Ke and J. Jin, "Special invited paper: The SCORE normalization, especially for heterogeneous network and text data," *Stat.*, vol. 12, no. 1, p. e545, Jan. 2023.
- [15] S. Dhara, J. Gaudio, E. Mossel, and C. Sandon, "The power of two matrices in spectral algorithms for community recovery," *IEEE Trans. Inf. Theory*, vol. 70, no. 5, pp. 3599–3621, May 2024.

- [16] J. Lei and A. Rinaldo, "Consistency of spectral clustering in stochastic block models," *Ann. Statist.*, vol. 43, no. 1, pp. 215–237, Feb. 2015.
- [17] T. Qin and K. Rohe, "Regularized spectral clustering under the degree-corrected stochastic blockmodel," in *Proc. Adv. Neural Inf. Process. Syst.*, 2013, pp. 3120–3128.
- [18] Z. Zhou and A. A. Amini, "Analysis of spectral clustering algorithms for community detection: The general bipartite setting," *J. Mach. Learn. Res.*, vol. 20, no. 47, pp. 1–47, 2019.
- [19] C. Gao, Z. Ma, A. Y. Zhang, and H. H. Zhou, "Achieving optimal misclassification proportion in stochastic block model," *J. Mach. Learn. Res.*, vol. 18, no. 60, pp. 1–45, 2017.
- [20] Y. Lu and H. H. Zhou, "Statistical and computational guarantees of Lloyd's algorithm and its variants," 2016, *arXiv:1612.02099*.
- [21] M. Xu, V. Jog, and P.-L. Loh, "Optimal rates for community estimation in the weighted stochastic block model," *Ann. Statist.*, vol. 48, no. 1, pp. 183–204, Feb. 2020.
- [22] R. Han, Y. Luo, M. Wang, and A. R. Zhang, "Exact clustering in tensor block model: Statistical optimality and computational limit," *J. Roy. Stat. Soc., Ser. B, Stat. Methodol.*, vol. 84, no. 5, pp. 1666–1698, Nov. 2022.
- [23] F. Gao, Z. Ma, and H. Yuan, "Community detection in sparse latent space models," *J. Mach. Learn. Res.*, vol. 23, no. 322, pp. 1–50, 2022.
- [24] E. Abbe, J. Fan, K. Wang, and Y. Zhong, "Entrywise eigenvector analysis of random matrices with low expected rank," *Ann. Statist.*, vol. 48, no. 3, p. 1452, Jun. 2020.
- [25] P. W. Holland, K. B. Laskey, and S. Leinhardt, "Stochastic blockmodels: First steps," *Social Netw.*, vol. 5, no. 2, pp. 109–137, Jun. 1983.
- [26] A. Y. Zhang and H. H. Zhou, "Minimax rates of community detection in stochastic block models," *Ann. Statist.*, vol. 44, no. 5, pp. 2252–2280, Oct. 2016.
- [27] M. Pensky and T. Zhang, "Spectral clustering in the dynamic stochastic block model," *Electron. J. Statist.*, vol. 13, no. 1, pp. 678–709, Jan. 2019.
- [28] A. Joseph and B. Yu, "Impact of regularization on spectral clustering," *Ann. Statist.*, vol. 44, no. 4, pp. 1765–1791, Aug. 2016.
- [29] C. M. Le, E. Levina, and R. Vershynin, "Concentration and regularization of random graphs," *Random Struct. Algorithms*, vol. 51, no. 3, pp. 538–561, Oct. 2017.
- [30] P. Chin, A. Rao, and V. Vu, "Stochastic block model and community detection in sparse graphs: A spectral algorithm with optimal rate of recovery," in *Proc. Conf. Learn. Theory*, 2015, pp. 391–423.
- [31] M. Löffler, A. Y. Zhang, and H. H. Zhou, "Optimality of spectral clustering in the Gaussian mixture model," *Ann. Statist.*, vol. 49, no. 5, pp. 2506–2530, Oct. 2021.
- [32] E. Abbe, J. Fan, and K. Wang, "An ℓ_p theory of PCA and spectral clustering," *Ann. Statist.*, vol. 50, no. 4, pp. 2359–2385, Aug. 2022.
- [33] A. Y. Zhang and H. H. Zhou, "Leave-one-out singular subspace perturbation analysis for spectral clustering," 2022, *arXiv:2205.14855*.
- [34] P. Chen, C. Gao, and A. Y. Zhang, "Partial recovery for top-k ranking: Optimality of MLE and SubOptimality of the spectral method," *Ann. Statist.*, vol. 50, no. 3, pp. 1618–1652, Jun. 2022.
- [35] Y. Chen, J. Fan, C. Ma, and K. Wang, "Spectral method and regularized MLE are both optimal for top- K ranking," *Ann. Statist.*, vol. 47, no. 4, p. 2204, Aug. 2019.
- [36] Y. Chen, Y. Chi, J. Fan, and C. Ma, "Spectral methods for data science: A statistical perspective," *Found. Trends Mach. Learn.*, vol. 14, no. 5, pp. 566–806, 2021.
- [37] M. E. J. Newman, "Modularity and community structure in networks," *Proc. Nat. Acad. Sci. USA*, vol. 103, no. 23, pp. 8577–8582, 2006.
- [38] C. E. Priebe et al., "On a two-truths phenomenon in spectral graph clustering," *Proc. Nat. Acad. Sci. USA*, vol. 116, no. 13, pp. 5995–6000, Mar. 2019.
- [39] P. Sarkar and P. J. Bickel, "Role of normalization in spectral clustering for stochastic blockmodels," *Ann. Statist.*, vol. 43, no. 3, pp. 962–990, Jun. 2015.
- [40] J. Cape, M. Tang, and C. E. Priebe, "On spectral embedding performance and elucidating network structure in stochastic blockmodel graphs," *Netw. Sci.*, vol. 7, no. 3, pp. 269–291, Sep. 2019.
- [41] C. Gao, Z. Ma, A. Y. Zhang, and H. H. Zhou, "Community detection in degree-corrected block models," *Ann. Statist.*, vol. 46, no. 5, p. 2153, Oct. 2018.

Anderson Ye Zhang received the B.S. degree in statistics from Zhejiang University, Hangzhou, China, in 2013, and the Ph.D. degree in statistics and data science from Yale University in 2018. He was a William H. Kruskal Instructor with the Department of Statistics, The University of Chicago, in 2018. He has been an Assistant Professor of statistics and data science with the University of Pennsylvania, since 2019.