Supplementary Material to "Distributed Community Detection for Large Scale Networks Using Stochastic Block Model"

APPENDIX A

Appendix A.1: Notations, Useful Lemmas and Propositions

We first define several notations which will be used in the rest of the proof. Let

$$\widetilde{L}^{(\mathcal{S}_m)} = \begin{pmatrix} \mathbf{0}_{\overline{n}_m \times \overline{n}_m} & L^{(\mathcal{S}_m)} \\ L^{(\mathcal{S}_m)\top} & \mathbf{0}_{l \times l} \end{pmatrix}, \quad \widetilde{\mathcal{L}}^{(\mathcal{S}_m)} = \begin{pmatrix} \mathbf{0}_{\overline{n}_m \times \overline{n}_m} & \mathcal{L}^{(\mathcal{S}_m)} \\ \mathcal{L}^{(\mathcal{S}_m)\top} & \mathbf{0}_{l \times l} \end{pmatrix}.$$

Similarly, define $\widetilde{A}^{(\mathcal{S}_m)}, \widetilde{\mathcal{A}}^{(\mathcal{S}_m)} \in \mathbb{R}^{(n_m+2l)\times(n_m+2l)}$ in the same way. In addition, define $\widetilde{D}^{(\mathcal{S}_m)} = \mathrm{diag}\{D^{(\mathcal{S}_m)}, F^{(\mathcal{S}_m)}\}$ and $\widetilde{\mathcal{D}}^{(\mathcal{S}_m)} = \mathrm{diag}\{\mathcal{D}^{(\mathcal{S}_m)}, \mathcal{F}^{(\mathcal{S}_m)}\}$.

Proposition 1. *Assume the same conditions with Theorem* **??**. *Then we have*

$$\big\|\widetilde{L}^{(\mathcal{S}_m)} - \widetilde{\mathcal{L}}^{(\mathcal{S}_m)}\big\|_{\max} \le 4\sqrt{3}\sqrt{\frac{\log(4(n_m+2l)/\epsilon_m)}{\delta_m}}. \tag{A.1}$$

with probability at least $1-\epsilon_m$.

Proof. Note that $\widetilde{D}^{(\mathcal{S}_m)}$ and $\widetilde{L}^{(\mathcal{S}_m)}$ are dependent with each other. Therefore, we add an intermediate step by using the matrix $\widetilde{C}^{(\mathcal{S}_m)} \stackrel{\text{def}}{=} (\widetilde{\mathcal{D}}^{(\mathcal{S}_m)})^{-1/2}\widetilde{A}^{(\mathcal{S}_m)}(\widetilde{\mathcal{D}}^{(\mathcal{S}_m)})^{-1/2}$. Hence we have

$$\big\|\widetilde{L}^{(\mathcal{S}_m)} - \widetilde{\mathcal{L}}^{(\mathcal{S}_m)}\big\|_{\max} \le \big\|\widetilde{L}^{(\mathcal{S}_m)} - \widetilde{C}^{(\mathcal{S}_m)}\big\|_{\max} + \big\|\widetilde{\mathcal{L}}^{(\mathcal{S}_m)} - \widetilde{C}^{(\mathcal{S}_m)}\big\|_{\max}. \tag{A.2}$$

Define $\nu = \sqrt{3\log(4(n_m+2l)/\epsilon_m)/\delta_m}$. Then we have $\nu \le 1$ for sufficiently large l. By (A.3) and (A.4) of Proposition 2, we have $\|\widetilde{L}^{(\mathcal{S}_m)} - \widetilde{\mathcal{L}}^{(\mathcal{S}_m)}\|_{\max} \le \nu^2 + 3\nu \le 4\nu$. This yields (A.1).

$\square$

Proposition 2. *Assume the same conditions in Theorem ??. Let* $\widetilde{C}^{(\mathcal{S}_m)} \overset{\text{def}}{=} (\widetilde{\mathcal{D}}^{(\mathcal{S}_m)})^{-1/2}\widetilde{A}^{(\mathcal{S}_m)}$ $(\widetilde{\mathcal{D}}^{(\mathcal{S}_m)})^{-1/2}$. *Then we have with probability at least* $1-\epsilon_m/2$

$$\big\|\widetilde{L}^{(\mathcal{S}_m)} - \widetilde{C}^{(\mathcal{S}_m)}\big\|_{\max} \le \nu^2 + 2\nu, \tag{A.3}$$

$$\big\|\widetilde{\mathcal{L}}^{(\mathcal{S}_m)} - \widetilde{C}^{(\mathcal{S}_m)}\big\|_{\max} \le \nu, \tag{A.4}$$

where

$$\nu = \sqrt{3\log\{4(n_m+2l)/\epsilon_m\}/\delta_m}. \tag{A.5}$$

Proof. Note that $\widetilde{\mathcal{D}}_{ii}^{(\mathcal{S}_m)} \ge \delta_m$ for $i = 1, ..., \overline{n}_m + l$. We prove (A.3) and (A.4) respectively as follows.

1. Proof of (A.3).

It can be derived that

$$\begin{aligned}
&\big\|\widetilde{L}^{(\mathcal{S}_m)} - \widetilde{C}^{(\mathcal{S}_m)}\big\|_{\max} = \big\|\widetilde{L}^{(\mathcal{S}_m)} - (\widetilde{\mathcal{D}}^{(\mathcal{S}_m)})^{-1/2}(\widetilde{D}^{(\mathcal{S}_m)})^{1/2}\widetilde{L}^{(\mathcal{S}_m)}(\widetilde{D}^{(\mathcal{S}_m)})^{1/2}(\widetilde{\mathcal{D}}^{(\mathcal{S}_m)})^{-1/2}\big\|_{\max} \\
&\le \Big\|\big\{I - (\widetilde{\mathcal{D}}^{(\mathcal{S}_m)})^{-1/2}(\widetilde{D}^{(\mathcal{S}_m)})^{1/2}\big\}\widetilde{L}^{(\mathcal{S}_m)}(\widetilde{D}^{(\mathcal{S}_m)})^{1/2}(\widetilde{\mathcal{D}}^{(\mathcal{S}_m)})^{-1/2}\Big\|_{\max} \\
&+ \Big\|\widetilde{L}^{(\mathcal{S}_m)}\big\{I - (\widetilde{D}^{(\mathcal{S}_m)})^{1/2}(\widetilde{\mathcal{D}}^{(\mathcal{S}_m)})^{-1/2}\big\}\Big\|_{\max} \overset{\text{def}}{=} \Delta_1 + \Delta_2.
\end{aligned}$$

We then deal with Δ_1 and Δ_2 respectively. Note that we have $\|\widetilde{L}^{(\mathcal{S}_m)}\|_{\max} \le 1$ and $\widetilde{D}^{(\mathcal{S}_m)}$, $\widetilde{\mathcal{D}}^{(\mathcal{S}_m)}$ are diagonal matrices. Hence we have $\Delta_2 \le \big\|I - (\widetilde{\mathcal{D}}^{(\mathcal{S}_m)})^{-1/2}(\widetilde{D}^{(\mathcal{S}_m)})^{1/2}\big\|_{\max}$

and

$$\Delta_1 \le \big\|I - (\widetilde{\mathcal{D}}^{(\mathcal{S}_m)})^{-1/2}(\widetilde{D}^{(\mathcal{S}_m)})^{1/2}\big\|_{\max}\big\|(\widetilde{\mathcal{D}}^{(\mathcal{S}_m)})^{-1/2}(\widetilde{D}^{(\mathcal{S}_m)})^{1/2}\big\|_{\max}$$
$$\le \big\|I - (\widetilde{\mathcal{D}}^{(\mathcal{S}_m)})^{-1/2}(\widetilde{D}^{(\mathcal{S}_m)})^{1/2}\big\|_{\max}^2 + \big\|I - (\widetilde{\mathcal{D}}^{(\mathcal{S}_m)})^{-1/2}(\widetilde{D}^{(\mathcal{S}_m)})^{1/2}\big\|_{\max}.$$

Then it suffices to bound $\big\|I - (\widetilde{\mathcal{D}}^{(\mathcal{S}_m)})^{-1/2}(\widetilde{D}^{(\mathcal{S}_m)})^{1/2}\big\|_{\max}$. By **?**, it holds that

$$\mathbb{P}\Big(\big|\widetilde{D}_{ii}^{(\mathcal{S}_m)} - \widetilde{\mathcal{D}}_{ii}^{(\mathcal{S}_m)}\big| \ge \lambda\Big) \le \exp\Big(-\frac{\lambda^2}{2\widetilde{\mathcal{D}}_{ii}^{(\mathcal{S}_m)}}\Big) + \exp\Big(-\frac{\lambda^2}{2\widetilde{\mathcal{D}}_{ii}^{(\mathcal{S}_m)} + 2\lambda/3}\Big). \tag{A.6}$$

Note that,

$$\Big\|(\widetilde{\mathcal{D}}^{(\mathcal{S}_m)})^{-1/2}(\tilde{D}^{(\mathcal{S}_m)})^{1/2} - I\Big\|_{\max} = \max_i \big|(\widetilde{D}_{ii}^{(\mathcal{S}_m)})^{1/2}(\widetilde{\mathcal{D}}_{ii}^{(\mathcal{S}_m)})^{-1/2} - 1\big|$$
$$\le \max_i \big|\widetilde{D}_{ii}^{(\mathcal{S}_m)}/\widetilde{\mathcal{D}}_{ii}^{(\mathcal{S}_m)} - 1\big|$$

This implies for any fixed ν,

$$\mathbb{P}\Big(\big\|(\widetilde{\mathcal{D}}^{(\mathcal{S}_m)})^{-1/2}(\widetilde{D}^{(\mathcal{S}_m)})^{1/2} - I\big\|_{\max} \ge \nu\Big) \le \mathbb{P}\Big(\max_i \big|\widetilde{D}_{ii}^{(\mathcal{S}_m)}/\widetilde{\mathcal{D}}_{ii}^{(\mathcal{S}_m)} - 1\big| \ge \nu\Big)$$
$$\le \sum_i \mathbb{P}\Big\{\Big|\widetilde{D}_{ii}^{(\mathcal{S}_m)} - \widetilde{\mathcal{D}}_{ii}^{(\mathcal{S}_m)}\Big| \ge \nu\widetilde{\mathcal{D}}_{ii}^{(\mathcal{S}_m)}\Big\}$$

By using (A.6) we further have

$$\mathbb{P}\Big\{\Big|\widetilde{D}_{ii}^{(\mathcal{S}_m)} - \widetilde{\mathcal{D}}_{ii}^{(\mathcal{S}_m)}\Big| \ge \nu\widetilde{\mathcal{D}}_{ii}^{(\mathcal{S}_m)}\Big\} \le \exp\big(-\nu^2\widetilde{\mathcal{D}}_{ii}^{(\mathcal{S}_m)}/2\big) + \exp\big\{-\nu^2\widetilde{\mathcal{D}}_{ii}^{(\mathcal{S}_m)}/(2+2\nu/3)\big\}$$
$$\le 2\exp\big(-\nu^2\widetilde{\mathcal{D}}_{ii}^{(\mathcal{S}_m)}/3\big)$$

Taking ν as in (A.5), it can be verified that

$$\exp\big(-\nu^2\widetilde{\mathcal{D}}_{ii}^{(\mathcal{S}_m)}/3\big) \le 2\exp\{-\log(4(n_m+2l)/\epsilon_m)\} = \epsilon_m/\{2(n_m+2l)\}.$$

Consequently (A.3) holds.

2. Proof of (A.4).

We bound the second part using the following concentration inequality given by **?**.

Lemma 1. *Let $X_1, X_2, \dots, X_m$ be independent random $N \times N$ Hermitian matrices. Moreover, assume that $\|X_i - \mathbb{E}(X_i)\|_{\max} \le M$ for all i, and $c^2 = \|\sum \operatorname{var}(X_i)\|_{\max}$. Let $X = \sum X_i$. Then for any $\nu > 0$*

$$\mathbb{P}(\|X - \mathbb{E}(X)\|_{\max} \ge \nu) \le 2N \exp\left(-\frac{\nu^2}{2c^2 + 2M\nu/3}\right)$$

Denote $E^{i,j} \in \mathbb{R}^{(n_m+2l)\times(n_m+2l)}$ with 1 in the $(i,j),(j,i)$ positions and 0 elsewhere, and define $X^{i,n_m+l+j} = (\mathcal{D}_{ii}^{(\mathcal{S}_m)}\mathcal{F}_{jj}^{(\mathcal{S}_m)})^{-1/2}(A_{ij}^{(\mathcal{S}_m)} - \mathcal{A}_{ij}^{(\mathcal{S}_m)})E^{i,n_m+l+j}, i = 1,\dots,n_m+l, j = 1,\dots,l$. Then we have

$$\|\widetilde{C}^{(\mathcal{S}_m)} - \widetilde{\mathcal{L}}^{(\mathcal{S}_m)}\|_{\max} = \Big\|\sum_{i=1}^{n_m+l}\sum_{j=1}^{l} X^{i,n_m+l+j}\Big\|_{\max},$$

due to that $\min_j \mathcal{F}_{jj}^{(\mathcal{S}_m)} \ge \min_i \mathcal{D}_{ii}^{(\mathcal{S}_m)}$. As a result, X^{i,n_m+l+j} are independent random Hermitian matrices. We then derive M and c^2 in this context and then the results can be obtained by Lemma 1. First note that $\mathbb{E}\left[X^{i,n_m+l+j}\right] = \mathbf{0}$. We then have

$$\|X^{i,n_m+l+j}\|_{\max} \le 1/\sqrt{\mathcal{D}_{ii}^{(\mathcal{S}_m)}\mathcal{F}_{jj}^{(\mathcal{S}_m)}} \le 1/\delta_m \stackrel{\text{def}}{=} M.$$

Next, note that $\mathbb{E}\big[\big(X^{i,\overline{n}_m+j}\big)^2\big] = \big(1/\mathcal{D}_{ii}^{(\mathcal{S}_m)}\mathcal{F}_{jj}^{(\mathcal{S}_m)}\big)\big[\mathcal{A}_{ij}^{(\mathcal{S}_m)}\big(1-\mathcal{A}_{ij}^{(\mathcal{S}_m)}\big)\big]\left(E^{ii}+E^{\overline{n}_m+j,\overline{n}_m+j}\right) \stackrel{\text{def}}{=}$

$v_{ij}(E^{ii} + E^{\overline{n}_m+j,\overline{n}_m+j})$. This leads to

$$\Big\| \sum_{i=1}^{n_m+l} \sum_{j=1}^{l} \mathbb{E}\left[\left(X^{i,n_m+l+j}\right)^2 \right] \Big\|_{\max} = \Big\| \sum_{i=1}^{n_m+l} \sum_{j=1}^{l} v_{ij}(E^{ii} + E^{n_m+l+j,n_m+l+j}) \Big\|_{\max}$$
$$= \max\Big\{ \max_{1\le i\le \overline{n}_m} \sum_j v_{ij}, \max_{1\le j\le l} \sum_i v_{ij} \Big\} \le \frac{1}{\delta_m} \overset{\text{def}}{=} c^2,$$

where the last inequality holds because

$$\sum_j v_{ij} \le \frac{1}{\delta_m} \sum_{j=1}^{l} \frac{\mathcal{A}_{ij}^{(\mathcal{S}_m)}}{\mathcal{D}_{ii}^{(\mathcal{S}_m)}} = \frac{1}{\delta_m}, \quad \sum_i v_{ij} \le \frac{1}{\delta_m} \sum_{i=1}^{\overline{n}_m} \frac{\mathcal{A}_{ij}^{(\mathcal{S}_m)}}{\mathcal{F}_{ii}^{(\mathcal{S}_m)}} = \frac{1}{\delta_m}.$$

By assumption $\delta_m > 3\log(n_m + 2l) + 3\log(4/\epsilon_m)$, we have $\nu < 1$. Applying Lemma 1, we have

$$\mathbb{P}(\|\widetilde{C}^{(\mathcal{S}_m)} - \widetilde{\mathcal{L}}^{(\mathcal{S}_m)}\|_{\max} \ge \nu) \le 2(n_m + 2l) \exp\left\{ -\frac{2\log(4(n_m+2l)/\epsilon_m)/\delta_m}{2/\delta_m + 2\nu/3\delta_m} \right\}$$
$$\le 2(n_m + 2l) \exp\Big\{ -\frac{3\log(4(n_m+2l)/\epsilon_m)}{3} \Big\} \le \epsilon/2$$

This completes the proof.

$\square$

Lemma 2. *Let $\lambda_{1,0} \ge \lambda_{2,0} \ge \cdots \ge \lambda_{K,0} > 0$ be the top K singular values of $\mathcal{L}_0$. Define $\delta_0 = \min_i \mathcal{D}_{0,ii}$. Then for any $\epsilon > 0$ and $\delta_0 > 3l\log(2l) + 3\log(4/\epsilon)$, with probability at least $1-\epsilon$ it holds*

$$\|\widehat{U}_0 - U_0 Q_0\|_F \le \frac{8\sqrt{6}}{\lambda_{K,0}} \sqrt{\frac{K\log(8l/\epsilon)}{\delta_0}} \tag{A.7}$$

where $Q_0 \in \mathbb{R}^{K\times K}$ is a $K \times K$ orthogonal matrix.

Proof. The proof follows the same procedure as in Theorem **??**. $\square$

Lemma 3. *Define $P_0 = \max_{j=1,\ldots,l} (\Theta_0^\top \Theta_0)_{jj}$. Denote $\mathcal{M}$ as the index set of misclustered* *nodes on the master server. Then for any ϵ and $\delta_0 > 3l\log(2l) + 3\log(4/\epsilon)$, it holds with probability $1-\epsilon$ that*

$$|\mathcal{M}| \leq \frac{3072 P_0 K \log(8l/\epsilon)}{\delta_0 \lambda_{K,0}^2}$$

Proof. Under the procedure in **?**, it could be verified that

$$|\mathcal{M}| \leq 8P_0 \|\widehat{U}_0 - U_0 Q_0\|_F^2. \tag{A.8}$$

Combining (A.7) and (A.8) yields the result. $\square$

APPENDIX B: Proof of Propositions

Appendix B.1: Proof of Proposition **??**

STEP 1. We first explore the spectral structure of $\mathcal{L}$ and $\mathcal{L}_0$. Construct a matrix $B_L \in \mathbb{R}^{K\times K}$ such that $\mathcal{L} = \Theta B_L \Theta^\top$. Define $D_B = \mathrm{diag}(B\Theta^\top \mathbf{1}_N) \in \mathbb{R}^{K\times K}$ where $\mathbf{1}_N$ is an $N \times 1$ vector with all entries 1. Denote Θ_i as the ith row of Θ. Note that for any i, j,

$$\mathcal{L}_{ij} = \frac{\mathcal{A}_{ij}}{\sqrt{\mathcal{D}_{ii}\mathcal{D}_{jj}}} = \Theta_i D_B^{-1/2} B D_B^{-1/2} \Theta_j^\top. \tag{B.1}$$

Consequently, define $B_L = D_B^{-1/2} B D_B^{-1/2}$. It follows $\mathcal{L} = \Theta B_L \Theta^\top$.

Similarly, For $\mathcal{L}_0$, define $D_{B_0} = \mathrm{diag}(B\Theta_0^\top \mathbf{1}_N) \in \mathbb{R}^{K\times K}$ and $B_{L_0} = D_{B_0}^{-1/2} B D_{B_0}^{-1/2}$, it can be obtained that $\mathcal{L}_0 = \Theta_0 B_{L_0} \Theta_0^\top$.

Step 2. Denote $\Lambda = \Theta^\top\Theta$, $\Lambda_0 = \Theta_0^\top\Theta_0$. Construct $\mathcal{L}$ and $\mathcal{L}_0$ as

$$\mathcal{L} = \Theta\Lambda^{-1/2}\Lambda^{1/2}B_L\Lambda^{1/2}\Lambda^{-1/2}\Theta^\top,$$
$$\mathcal{L}_0 = \Theta_0\Lambda_0^{-1/2}\Lambda_0^{1/2}B_{L_0}\Lambda_0^{1/2}\Lambda_0^{-1/2}\Theta_0^\top.$$

Conduct eigen-decompositions as $\Lambda^{1/2}B_L\Lambda^{1/2} = \mu U\mu^\top$ and $\Lambda_0^{1/2}B_{L_0}\Lambda_0^{1/2} = \mu_0 U_0\mu_0^\top$, where μ,$\mu_0 \in \mathbb{R}^{K\times K}$ are orthogonal matrices and U,$U_0 \in \mathbb{R}^{K\times K}$ are diagonal matrices. By the assumption $m_{0k}/m_k = l/N = r_0$, we have $\Lambda_0 = r_0\Lambda$ and $\Theta_0^\top\mathbf{1}_l = r_0\Theta^\top\mathbf{1}_N$.

Step 3. Recall the eigen-decomposition of $\mathcal{L}_0$ and $\mathcal{L}$, by Step 2, we know that $\Lambda^{1/2}B_L\Lambda^{1/2}$ and $\Lambda_0^{1/2}B_{L_0}\Lambda_0^{1/2}$ differ from a scalar multiplication, thus $\mu = \mu_0$. Subsequently, $\mathcal{L}$ and $\mathcal{L}_0$ have the following eigen-decomposition:

$$\mathcal{L} = \Theta\Lambda^{-1/2}\mu U\mu^\top\Lambda^{-1/2}\Theta^\top,$$
$$\mathcal{L}_0 = \Theta_0\Lambda_0^{-1/2}\mu_0 U_0\mu_0^\top\Lambda_0^{-1/2}\Theta_0^\top.$$

Further note that $U^{(K)} = \Lambda^{-1/2}\mu$ and $U_0^{(K)} = \Lambda_0^{-1/2}\mu$, then the result naturally holds.

Appendix B.2: Proof of Proposition **??**

Proof. We separate the proof into two steps.

In the first step, we show that $\mathcal{L}^{(\mathcal{S}_m)}$ can be expressed as

$$\mathcal{L}^{(\mathcal{S}_m)} = \Theta^{(\mathcal{S}_m)}(\mathcal{D}_B^{(\mathcal{S}_m)})^{-1/2}B(\mathcal{F}_B^{(\mathcal{S}_m)})^{-1/2}\Theta_0^\top, \tag{B.2}$$

where $\mathcal{D}_B^{(\mathcal{S}_m)} = \mathrm{diag}\{B\Theta_0^\top\mathbf{1}_l\} \in \mathbb{R}^{K\times K}$ and $\mathcal{F}_B^{(\mathcal{S}_m)} = \mathrm{diag}\{B(\Theta^{(\mathcal{S}_m)})^\top\mathbf{1}_{\overline{n}_m}\} \in \mathbb{R}^{K\times K}$. In the second step, based on the form in (B.2), we show that $U^{(\mathcal{S}_m)} = \Theta^{(\mathcal{S}_m)}\mu$ is the eigenvector matrix of $\mathcal{L}^{(\mathcal{S}_m)}\mathcal{L}^{(\mathcal{S}_m)\top}$ and μ is a full rank matrix. This leads to the final result.

Step 1. Note that $\mathcal{A}^{(\mathcal{S}_m)}\mathbf{1}_l = \Theta^{(\mathcal{S}_m)}B\Theta_0^\top\mathbf{1}_l$ and $\mathcal{A}^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m} = \Theta_0 B\Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m}$. Therefore, we have $\mathcal{D}^{(\mathcal{S}_m)} = \mathrm{diag}\{\mathcal{A}^{(\mathcal{S}_m)}\mathbf{1}_l\}$ and $\mathcal{F}^{(\mathcal{S}_m)} = \mathrm{diag}\{\mathcal{A}^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m}\}$. Then we have

$$\begin{aligned}
\mathcal{D}_{ii}^{(\mathcal{S}_m)} &= \Theta_i^{(\mathcal{S}_m)}B\Theta_0^\top\mathbf{1}_l = B_{g_i}^\top\Theta_0^\top\mathbf{1}_l \\
\mathcal{F}_{ii}^{(\mathcal{S}_m)} &= \Theta_{0i}B\Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m} = B_{g_i}^\top\Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m}.
\end{aligned}$$

Then it can be obtained that

$$\begin{aligned}
\mathcal{L}_{ij}^{(\mathcal{S}_m)} &= \frac{\mathcal{A}_{ij}^{(\mathcal{S}_m)}}{\sqrt{\mathcal{D}_{ii}^{(\mathcal{S}_m)}\mathcal{F}_{jj}^{(\mathcal{S}_m)}}} = \big(B_{g_i}^\top\Theta_0^\top\mathbf{1}_l\big)^{-1/2}\big(\Theta_i^{(\mathcal{S}_m)\top}B\Theta_{0j}\big)\big(B_{g_i}^\top\Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m}\big)^{-1/2} \\
&= \mathbf{e}_{g_i}^\top(\mathcal{D}_B^{(\mathcal{S}_m)})^{-1/2}\big(\Theta_i^{(\mathcal{S}_m)\top}B\Theta_{0j}\big)(\mathcal{F}_B^{(\mathcal{S}_m)})^{-1/2}\mathbf{e}_{g_j} \\
&= \Theta_i^{(\mathcal{S}_m)\top}(\mathcal{D}_B^{(\mathcal{S}_m)})^{-1/2}B(\mathcal{F}_B^{(\mathcal{S}_m)})^{-1/2}\Theta_{0j}.
\end{aligned}$$

This immediately yields (B.2).

Step 2. In the following we show that the eigen-decomposition of $\mathcal{L}^{(\mathcal{S}_m)}(\mathcal{L}^{(\mathcal{S}_m)})^\top$ takes the form

$$\mathcal{L}^{(\mathcal{S}_m)}(\mathcal{L}^{(\mathcal{S}_m)})^\top = (\Theta^{(\mathcal{S}_m)}\mu)\Lambda(\Theta^{(\mathcal{S}_m)}\mu)^\top,$$

where $U^{(\mathcal{S}_m)} = \Theta^{(\mathcal{S}_m)}\mu \in \mathbb{R}^{\overline{n}_m\times K}$ is the eigenvector matrix and $\Lambda \in \mathbb{R}^{K\times K}$ is the diagonal eigenvalue matrix. To this end, we first write

$$\mathcal{L}^{(\mathcal{S}_m)}(\mathcal{L}^{(\mathcal{S}_m)})^\top = \Theta^{(\mathcal{S}_m)}(\mathcal{D}_B^{(\mathcal{S}_m)})^{-1/2}B(\mathcal{F}_B^{(\mathcal{S}_m)})^{-1/2}\Theta_0^\top\Theta_0(\mathcal{F}_B^{(\mathcal{S}_m)})^{-1/2}B(\mathcal{D}_B^{(\mathcal{S}_m)})^{-1/2}(\Theta^{(\mathcal{S}_m)})^\top,$$

$\stackrel{\mathrm{def}}{=} \Theta^{(\mathcal{S}_m)}B_L(\Theta^{(\mathcal{S}_m)})^\top$. Define $\Delta = (\Theta^{(\mathcal{S}_m)})^\top\Theta^{(\mathcal{S}_m)}$. Then conduct the following eigen-decomposition as $\Delta^{1/2}B_L\Delta^{1/2} = V\Lambda V^\top$. This further implies

$$\Theta^{(\mathcal{S}_m)}B_L(\Theta^{(\mathcal{S}_m)})^\top = (\Theta^{(\mathcal{S}_m)}\Delta^{-1/2})\Delta^{1/2}B_L\Delta^{1/2}(\Delta^{-1/2}\Theta^{(\mathcal{S}_m)\top})$$
$$= (\Theta^{(\mathcal{S}_m)}\Delta^{-1/2})V\Lambda V^\top(\Delta^{-1/2}\Theta^{(\mathcal{S}_m)\top}) \stackrel{\text{def}}{=} (\Theta^{(\mathcal{S}_m)}\mu)\Lambda(\Theta^{(\mathcal{S}_m)}\mu)^\top.$$

Note that $(\Theta^{(\mathcal{S}_m)}\mu)^\top(\Theta^{(\mathcal{S}_m)}\mu) = I_K$. By the uniqueness of the eigen-decomposition, we know $U^{(\mathcal{S}_m)} = \Theta^{(\mathcal{S}_m)}\mu$ is the eigenvector matrix of $\mathcal{L}^{(\mathcal{S}_m)}(\mathcal{L}^{(\mathcal{S}_m)})^\top$. Further note that the matrix μ is full rank, then we can conclude that

$$\Theta_i^{(\mathcal{S}_m)}\mu = \Theta_j^{(\mathcal{S}_m)}\mu \Leftrightarrow \Theta_i^{(\mathcal{S}_m)} = \Theta_j^{(\mathcal{S}_m)}.$$

□

Appendix B.3: Proof of Proposition **??**

Proof. Denote $\mathcal{A}^{(\mathcal{S}_m\star)} = \Theta^{(\mathcal{S}_m)}B\Theta^{(\mathcal{S}_m)\top}$ and $\mathcal{L}^{(\mathcal{S}_m\star)}$ to be its Laplacian matrix, $U^{(\mathcal{S}_m\star)}$ be the K leading eigenvectors of $\mathcal{L}^{(\mathcal{S}_m\star)}$. We then have

$$\|U^{(\mathcal{S}_m)} - r_mU_mQ_m\|_F \le \|U^{\mathcal{S}_m} - U^{(\mathcal{S}_m\star)}Q_{m1}\|_F + \|U^{(\mathcal{S}_m\star)}Q_{m1} - r_mU_mQ_m\|_F,$$

where Q_{m1} is another $K\times K$ orthogonal matrix. In the following we show that

$$\|U^{(\mathcal{S}_m)} - U^{(\mathcal{S}_m\star)}Q_{m1}\|_F \le \frac{8\sqrt{2}K^2u_0u_m^2\max\{u_0^{1/2}, u_m^{1/2}\}\alpha^{(\mathcal{S}_m)1/2}}{\sigma_{\min}(B)^2b_{\min}^3d_0^2d_m^3}, \tag{B.3}$$
$$\|U^{(\mathcal{S}_m\star)}Q_{m1} - r_mU_mQ_m\|_F \le \frac{6\sqrt{2}Ku_m\max\{u_0^{1/2}, u_m^{1/2}\}\alpha^{(\mathcal{S}_m)1/2}}{\sigma_{\min}(B)b_{\min}^2d_0d_m^2(d_0+d_m)} + \frac{\alpha^{(\mathcal{S}_m)}}{d_0} \tag{B.4}$$

where Q_{m1} is a $K\times K$ orthogonal matrix. Then combining (B.3) and (B.4) yields (**??**). The proof is separated into three parts as follows.

0. Re-express $\mathcal{L}^{(\mathcal{S}_m)}$.

Firstly, we show that $\mathcal{L}^{(\mathcal{S}_m)}$ can be expressed as

$$\mathcal{L}^{(\mathcal{S}_m)} = \Theta^{(\mathcal{S}_m)}(\mathcal{D}_B^{(\mathcal{S}_m)})^{-1/2} B (\mathcal{F}_B^{(\mathcal{S}_m)})^{-1/2}\Theta_0^\top, \tag{B.5}$$

where $\mathcal{D}_B^{(\mathcal{S}_m)} = \mathrm{diag}\{B\Theta_0^\top \mathbf{1}_l\} \in \mathbb{R}^{K\times K}$ and $\mathcal{F}_B^{(\mathcal{S}_m)} = \mathrm{diag}\{B(\Theta^{(\mathcal{S}_m)})^\top \mathbf{1}_{\overline{n}_m}\} \in \mathbb{R}^{K\times K}$. Note that $\mathcal{A}^{(\mathcal{S}_m)}\mathbf{1}_l = \Theta^{(\mathcal{S}_m)}B\Theta_0^\top\mathbf{1}_l$ and $\mathcal{A}^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m} = \Theta_0 B\Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m}$. Therefore, we have $\mathcal{D}^{(\mathcal{S}_m)} = \mathrm{diag}\{\mathcal{A}^{(\mathcal{S}_m)}\mathbf{1}_l\}$ and $\mathcal{F}^{(\mathcal{S}_m)} = \mathrm{diag}\{\mathcal{A}^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m}\}$. Then we have

$$\begin{aligned}
\mathcal{D}_{ii}^{(\mathcal{S}_m)} &= \Theta_i^{(\mathcal{S}_m)\top} B\Theta_0^\top \mathbf{1}_l = B_{g_i}^\top \Theta_0^\top \mathbf{1}_l \\
\mathcal{F}_{ii}^{(\mathcal{S}_m)} &= \Theta_{0i}^\top B\Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m} = B_{g_i}^\top \Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m}.
\end{aligned}$$

Then it can be obtained that

$$\begin{aligned}
\mathcal{L}_{ij}^{(\mathcal{S}_m)} &= \frac{\mathcal{A}_{ij}^{(\mathcal{S}_m)}}{\sqrt{\mathcal{D}_{ii}^{(\mathcal{S}_m)}\mathcal{F}_{jj}^{(\mathcal{S}_m)}}} = \big(B_{g_i}^\top\Theta_0^\top\mathbf{1}_l\big)^{-1/2}\big(\Theta_i^{(\mathcal{S}_m)\top}B\Theta_{0j}\big)\big(B_{g_i}^\top\Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m}\big)^{-1/2} \\
&= \mathbf{e}_{g_i}^\top(\mathcal{D}_B^{(\mathcal{S}_m)})^{-1/2}\mathbf{e}_{g_i}\big(\Theta_i^{(\mathcal{S}_m)\top}B\Theta_{0j}\big)\mathbf{e}_{g_j}^\top(\mathcal{F}_B^{(\mathcal{S}_m)})^{-1/2}\mathbf{e}_{g_j} \\
&= \Theta_i^{(\mathcal{S}_m)\top}(\mathcal{D}_B^{(\mathcal{S}_m)})^{-1/2}B(\mathcal{F}_B^{(\mathcal{S}_m)})^{-1/2}\Theta_{0j},
\end{aligned}$$

This immediately yields (B.5). Similarly define $\mathcal{D}_B = \mathrm{diag}\{B\Theta^\top\mathbf{1}_N\}$. We have

$$\mathcal{L}^{(\mathcal{S}_m\star)} = \Theta^{(\mathcal{S}_m)}(\mathcal{F}_B^{(\mathcal{S}_m)})^{-1/2}B(\mathcal{F}_B^{(\mathcal{S}_m)})^{-1/2}\Theta^{(\mathcal{S}_m)\top} \tag{B.6}$$

$$\mathcal{L} = \Theta\mathcal{D}_B^{-1/2}B\mathcal{D}_B^{-1/2}\Theta^\top \tag{B.7}$$

Now we prove (B.3) and (B.4) respectively.

1. Proof of (B.3).

Denote $B^{(\mathcal{S}_m\star)} = (\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{1/2}(\mathcal{F}_B^{(\mathcal{S}_m)})^{-1/2}B(\mathcal{F}_B^{(\mathcal{S}_m)})^{-1/2}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{1/2}$, $B^{(\mathcal{S}_m)} = (\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{1/2}(\mathcal{D}_B^{(\mathcal{S}_m)})^{-1/2}B\ (\mathcal{F}_B^{(\mathcal{S}_m)})^{-1/2}(\Theta_0^\top\Theta_0)^{1/2}$. It is easy to verify that

$$\mathcal{L}^{(\mathcal{S}_m\star)}\mathcal{L}^{(\mathcal{S}_m\star)\top} = \Theta^{(\mathcal{S}_m)}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{-1/2}B^{(\mathcal{S}_m\star)}B^{(\mathcal{S}_m\star)\top}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{-1/2}\Theta^{(\mathcal{S}_m)\top}$$

$$\mathcal{L}^{(\mathcal{S}_m)}\mathcal{L}^{(\mathcal{S}_m)\top} = \Theta^{(\mathcal{S}_m)}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{-1/2}B^{(\mathcal{S}_m)}B^{(\mathcal{S}_m)\top}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{-1/2}\Theta^{(\mathcal{S}_m)\top}.$$

We separate the proof in following three steps.

Step 1.1 (Relate $\|U^{(\mathcal{S}_m)} - U^{(\mathcal{S}_m\star)}Q_{m1}\|_F$ to $\left\|B^{(\mathcal{S}_m\star)}B^{(\mathcal{S}_m\star)\top} - B^{(\mathcal{S}_m)}B^{(\mathcal{S}_m)\top}\right\|_{\max}$).

Denote $\mu^{(\mathcal{S}_m\star)}, \mu^{(\mathcal{S}_m)} \in \mathbb{R}^{K\times K}$ as the eigenvectors of $B^{(\mathcal{S}_m\star)}B^{(\mathcal{S}_m\star)\top}$ and $B^{(\mathcal{S}_m)}B^{(\mathcal{S}_m)\top}$, respectively. Then immediately we have $U^{(\mathcal{S}_m)} = \Theta^{(\mathcal{S}_m)}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{-1/2}\mu^{(\mathcal{S}_m)}$, $U^{(\mathcal{S}_m\star)} = \Theta^{(\mathcal{S}_m)}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{-1/2}\mu^{(\mathcal{S}_m\star)}$. Using Lemma 5.1 of **?**, we have

$$\|\mu^{(\mathcal{S}_m)} - \mu^{(\mathcal{S}_m\star)}Q_{m1}\|_F \leq \frac{2\sqrt{2}K}{\gamma_m}\left\|B^{(\mathcal{S}_m\star)}B^{(\mathcal{S}_m\star)\top} - B^{(\mathcal{S}_m)}B^{(\mathcal{S}_m)\top}\right\|_{\max},$$

where γ_m is the smallest eigenvalue of $B^{(\mathcal{S}_m)}B^{(\mathcal{S}_m)\top}$. Then

$$\begin{aligned}
&\|U^{(\mathcal{S}_m)} - U^{(\mathcal{S}_m\star)}Q_{m1}\|_F = \|\Theta^{(\mathcal{S}_m)}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{-1/2}(\mu^{(\mathcal{S}_m)} - \mu^{(\mathcal{S}_m\star)}Q_{m1})\|_F \\
&\leq \sigma_{\max}\big\{\Theta^{(\mathcal{S}_m)}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{-1/2}\big\}\|\mu^{(\mathcal{S}_m)} - \mu^{(\mathcal{S}_m\star)}Q_{m1}\|_F \\
&\leq \frac{2\sqrt{2}K}{\gamma_m}\left\|B^{(\mathcal{S}_m\star)}B^{(\mathcal{S}_m\star)\top} - B^{(\mathcal{S}_m)}B^{(\mathcal{S}_m)\top}\right\|_{\max}
\end{aligned}$$

where the last inequality is due to $\sigma_{\max}\big\{\Theta^{(\mathcal{S}_m)}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{-1/2}\big\} = 1$. In Step 1.2 and 1.3 we derive upper bound for $\left\|B^{(\mathcal{S}_m\star)}B^{(\mathcal{S}_m\star)\top} - B^{(\mathcal{S}_m)}B^{(\mathcal{S}_m)\top}\right\|_{\max}$ and lower bound for γ_m respectively.

Step 1.2 (Upper bound for $\left\|B^{(\mathcal{S}_m\star)}B^{(\mathcal{S}_m\star)\top} - B^{(\mathcal{S}_m)}B^{(\mathcal{S}_m)\top}\right\|_{\max}$).

Note here $\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)}$ and $\Theta_0^\top\Theta_0$ are diagonal matrices. Denote $o_i = (\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})_{ii}^{1/2}$ and $p_j = (\Theta_0^\top\Theta_0)_{jj}^{1/2}$, $i, j = 1, ..., K$. Then we have

$$B_{ij}^{(\mathcal{S}_m\star)} = \frac{o_i B_{ij} o_j}{\sqrt{(B_i^\top\Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m})(B_j^\top\Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m})}}$$
$$B_{ij}^{(\mathcal{S}_m)} = \frac{o_i B_{ij} p_j}{\sqrt{(B_i^\top\Theta_0^\top\mathbf{1}_l)(B_j^\top\Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m})}}$$

For convenience, denote $a_i = B_i^\top\Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m}$, $b_i = B_i^\top\Theta_0^\top\mathbf{1}_l$, $c_i = lB_i^\top\Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m}/\overline{n}_m$, $q_i = o_i\sqrt{l}/\sqrt{\overline{n}_m}$ then

$$\begin{aligned}
\left|\left(B^{(\mathcal{S}_m\star)}B^{(\mathcal{S}_m\star)\top} - B^{(\mathcal{S}_m)}B^{(\mathcal{S}_m)\top}\right)_{ij}\right| &= \Big|\sum_{k=1}^K \Big(\frac{o_io_jB_{ik}B_{jk}o_k^2}{\sqrt{a_ia_j}a_k} - \frac{o_io_jB_{ik}B_{jk}p_k^2}{\sqrt{b_ib_j}a_k}\Big)\Big| \\
&\le \frac{o_io_j}{a_k}\sum_{k=1}^K\Big|\frac{o_k^2}{\sqrt{a_ia_j}} - \frac{p_k^2}{\sqrt{b_ib_j}}\Big| = \frac{o_io_j}{a_k}\sum_{k=1}^K\Big|\frac{q_k^2}{\sqrt{c_ic_j}} - \frac{p_k^2}{\sqrt{b_ib_j}}\Big| \\
&= \frac{o_io_j}{a_k}\sum_{k=1}^K\Big|\frac{q_k^2\sqrt{b_ib_j} - p_k^2\sqrt{c_ic_j}}{\sqrt{c_ic_jb_ib_j}}\Big| \\
&\le \frac{o_io_j}{a_k}\sum_{k=1}^K\Big(\frac{\left|q_k^2 - p_k^2\right|}{\sqrt{c_ic_j}} + \frac{p_k^2|\sqrt{b_ib_j} - \sqrt{c_ic_j}|}{\sqrt{c_ic_jb_ib_j}}\Big)
\end{aligned}$$

We then give upper bounds for the two parts respectively as follows. First note that $a_k = B_k^\top\Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m} \ge Kb_{\min}\overline{n}_md_m$, $c_i = lB_i^\top\Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m}/\overline{n}_m \ge lKb_{\min}d_m\overline{n}_m/\overline{n}_m = lKb_{\min}d_m$, and $\left|q_k^2 - p_k^2\right| = \left|\left(l/\overline{n}_m\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)} - \Theta_0^\top\Theta_0\right)_{kk}\right| \le l\alpha^{(\mathcal{S}_m)}$. This leads to

$$\frac{o_io_j}{a_k}\sum_k\frac{\left|q_k^2 - p_k^2\right|}{\sqrt{c_ic_j}} \le \frac{u_m\alpha^{(\mathcal{S}_m)}}{Kb_{\min}^2d_m^2} \tag{B.8}$$

Next, for the second part we have $b_i = B_i^\top\Theta_0^\top\mathbf{1}_l \ge Klb_{\min}d_0$, $p_i^2 \le lu_0$. Next we discuss the upper bound for $|\sqrt{b_ib_j} - \sqrt{c_ic_j}|$. If $b_ib_j \ge c_ic_j$, then we have $\sqrt{b_ib_j} - \sqrt{c_ic_j} = \sqrt{(b_i - c_i + c_i)(b_j - c_j + c_j)} - \sqrt{c_ic_j} \le \sqrt{c_j|b_i - c_i|} + \sqrt{c_i|b_j - c_j|} + \sqrt{|b_i - c_i||b_j - c_j|}$. Otherwise, the upper bound is given by $\sqrt{b_j|b_i - c_i|} + \sqrt{b_i|c_j - b_j|} + \sqrt{|b_i - c_i||b_j - c_j|}$.

Consequently we have

$$|\sqrt{b_i b_j} - \sqrt{c_i c_j}| \le 2 \max_i \{\sqrt{b_i}, \sqrt{c_i}\} \max_i \sqrt{|b_i - c_i|} + \max_i |b_i - c_j|. \tag{B.9}$$

Since $b_i \le Klu_0$, $c_i \le Klu_m$, and $|b_i - c_i| = l|B_i^\top(\Theta_0^\top \mathbf{1}_l/l - \Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m}/\overline{n}_m)| \le Kl\alpha^{(\mathcal{S}_m)}$. As a result, we have $|\sqrt{b_i b_j} - \sqrt{c_i c_j}| \le 2Kl \max\{u_0^{1/2}, u_m^{1/2}\}\alpha^{(\mathcal{S}_m)1/2} + Kl\alpha^{(\mathcal{S}_m)} \le 3Kl \max\{u_0^{1/2}, u_m^{1/2}\}\alpha^{(\mathcal{S}_m)1/2}$, where the inequality is due to that $\alpha^{(\mathcal{S}_m)} \le \max\{u_0, u_m\}$. As a consequence, the upper bound for the second part is

$$\frac{o_i o_j}{a_k} \sum_{k=1}^{K} \frac{p_k^2 |\sqrt{b_i b_j} - \sqrt{c_i c_j}|}{\sqrt{c_i c_j b_i b_j}} \le \frac{3u_0 u_m \max\{u_0^{1/2}, u_m^{1/2}\}\alpha^{(\mathcal{S}_m)1/2}}{K^2 b_{\min}^3 d_0 d_m^2}. \tag{B.10}$$

Combing the results from (B.8) and (B.10), we obtain that

$$\left\|B^{(\mathcal{S}_m\star)}B^{(\mathcal{S}_m\star)\top} - B^{(\mathcal{S}_m)}B^{(\mathcal{S}_m)\top}\right\|_{\max} \le \frac{4u_m \max\{u_0^{1/2}, u_m^{1/2}\}\alpha^{(\mathcal{S}_m)1/2}}{K b_{\min}^3 d_0 d_m^2}.$$

Step 1.3 (Lower bound on γ_m). Recall that γ_m is the smallest eigenvalue of $B^{(\mathcal{S}_m)}B^{(\mathcal{S}_m)\top}$. Here we have $B^{(\mathcal{S}_m)}B^{(\mathcal{S}_m)\top} = (\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{1/2}(\mathcal{D}_B^{(\mathcal{S}_m)})^{-1/2}B(\mathcal{F}_B^{(\mathcal{S}_m)})^{-1/2}$ $(\Theta_0^\top\Theta_0)(\mathcal{F}_B^{(\mathcal{S}_m)})^{-1/2}B(\mathcal{D}_B^{(\mathcal{S}_m)})^{-1/2}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{1/2}$. Specifically $\Theta_0^\top\Theta_0$, $\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)}$, $\mathcal{F}_B^{(\mathcal{S}_m)}$, and $\mathcal{D}_B^{(\mathcal{S}_m)}$ are all diagonal matrices. As a result, $\lambda_{\min}(\Theta_0^\top\Theta_0) \ge ld_0$, $\lambda_{\min}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)}) \ge \overline{n}_m d_m$, $\lambda_{\max}(\mathcal{F}_B^{(\mathcal{S}_m)}) \le K\overline{n}_m u_m$, and $\lambda_{\max}(\mathcal{D}_B^{(\mathcal{S}_m)}) \le Klu_0$. Therefore we have

$$\gamma_m \ge \sigma_{\min}(B)^2 \frac{\overline{n}_m l d_0 d_m}{K^2 l \overline{n}_m u_0 u_m} = \sigma_{\min}(B)^2 \frac{d_0 d_m}{K^2 u_0 u_m}.$$

This leads to the final result.

2. Proof of (B.4).

Denote $B_L^{(\mathcal{S}_m)} = (\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{1/2}(\mathcal{F}_B^{(\mathcal{S}_m)})^{-1/2}B(\mathcal{F}_B^{(\mathcal{S}_m)})^{-1/2}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{1/2}$ and $B_L =$

$(\Theta^\top\Theta)^{1/2}\mathcal{D}_B^{-1/2}B\mathcal{D}_B^{-1/2}(\Theta^\top\Theta)^{1/2}$. According to (B.6) and (B.7), we have

$$\mathcal{L}^{(\mathcal{S}_m\star)} = \Theta^{(\mathcal{S}_m)}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{-1/2}B_L^{(\mathcal{S}_m)}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{-1/2}\Theta^{(\mathcal{S}_m)\top}$$

$$\mathcal{L} = \Theta(\Theta^\top\Theta)^{-1/2}B_L(\Theta^\top\Theta)^{-1/2}\Theta^\top$$

Denote $f_i = (\Theta^\top\Theta)_{ii}^{1/2}$. Note that Here we can write

$$(B_L)_{ij} = \frac{f_i B_{ij} f_j}{\sqrt{B_i^\top\Theta^\top\mathbf{1}_N B_j^\top\Theta^\top\mathbf{1}_N}}$$

$$(B_L^{(\mathcal{S}_m)})_{ij} = \frac{o_i B_{ij} o_j}{\sqrt{B_i^\top\Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m} B_j^\top\Theta^{(\mathcal{S}_m)\top}\mathbf{1}_{\overline{n}_m}}}$$

where $f_i = (\Theta^\top\Theta)_{ii}^{1/2}$ and B_i is the ith column of B. In the following we prove the upper bound in three steps.

Step 2.1. (Relate $\|U^{(\mathcal{S}_m\star)}Q_{m1} - r_m U_m Q_m\|_F$ to $\|N/\overline{n}_m B_L - B_L^{(\mathcal{S}_m)}\|_{\max}$)

Denote $\xi, \xi^{(\mathcal{S}_m)} \in \mathbb{R}^{K\times K}$ as the eigenvectors of B_L and $B_L^{(\mathcal{S}_m)}$, respectively. Assume that the smallest eigenvalue of $B_L^{(\mathcal{S}_m)}$ is τ_m. Note that scalar multiplication does not change the spectrum, using Lemma 5.1 of **?**, we have

$$\|\xi^{(\mathcal{S}_m)} - \xi Q_{m2}\|_F \leq \frac{2\sqrt{2}K}{\tau_m}\|\frac{N}{\overline{n}_m}B_L - B_L^{(\mathcal{S}_m)}\|_{\max}$$

Note that $U^{(\mathcal{S}_m\star)} = \Theta^{(\mathcal{S}_m)}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{-1/2}\xi^{(\mathcal{S}_m)}$, $U_m = \Theta^{(\mathcal{S}_m)}(\Theta^\top\Theta)^{-1/2}\xi$. Similar to

Step 1.1 we have

$$
\begin{aligned}
&\left\|U^{(\mathcal{S}_m\star)} - r_m U_m Q_{m2}\right\|_F = \left\|\Theta^{(\mathcal{S}_m)}((\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{-1/2}\xi^{(\mathcal{S}_m)} - \sqrt{\frac{N}{\overline{n}_m}}(\Theta^\top\Theta)^{-1/2}\xi Q_{m2})\right\|_F \\
&\quad \le \sigma_{\max}\{\Theta^{(\mathcal{S}_m)}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{-1/2}\}\left\|\xi^{(\mathcal{S}_m)} - \sqrt{\frac{N}{\overline{n}_m}}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{1/2}(\Theta^\top\Theta)^{-1/2}\xi Q_{m2}\right\|_F \\
&\quad \le \left\|\xi^{(\mathcal{S}_m)} - \sqrt{\frac{N}{\overline{n}_m}}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{1/2}(\Theta^\top\Theta)^{-1/2}\xi Q_{m2}\right\|_F,
\end{aligned}
$$

where the last inequality is due to that $\sigma_{\max}\{\Theta^{(\mathcal{S}_m)}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{-1/2}\} \le 1$. Furthermore, it is upper bounded by

$$
\begin{aligned}
&\left\|\xi^{(\mathcal{S}_m)} - \sqrt{\frac{N}{\overline{n}_m}}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{1/2}(\Theta^\top\Theta)^{-1/2}\xi Q_{m2}\right\|_F \\
&\quad \le \|\xi^{(\mathcal{S}_m)} - \xi Q_{m2}\|_F + \|\xi Q_{m2} - N^{1/2}/\overline{n}_m{}^{1/2}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{1/2}(\Theta^\top\Theta)^{-1/2}\xi Q_{m2}\|_F \\
&\quad \le \|\xi^{(\mathcal{S}_m)} - \xi Q_{m2}\|_F + \sigma_{\max}\{I - N^{1/2}/\overline{n}_m{}^{1/2}(\Theta^{(\mathcal{S}_m)\top}\Theta^{(\mathcal{S}_m)})^{1/2}(\Theta^\top\Theta)^{-1/2}\}\|\xi Q_{m2}\|_F \\
&\quad \le \|\xi^{(\mathcal{S}_m)} - \xi Q_{m2}\|_F + \frac{\alpha^{(\mathcal{S}_m)}}{d_0}
\end{aligned}
$$

where the last inequality holds because

$$
\left|1 - \sqrt{\frac{No_i^2/\overline{n}_m}{f_i^2}}\right| \le \left|1 - \frac{No_i^2/\overline{n}_m}{f_i^2}\right| = \frac{\left|f_i^2 - No_i^2/\overline{n}_m\right|}{f_i^2} \le \frac{N\alpha^{(\mathcal{S}_m)}}{Nd_0} = \frac{\alpha^{(\mathcal{S}_m)}}{d_0}.
$$

With a simple rotation using Q_{m1}, we have

$$
\|U^{(\mathcal{S}_m)\star}Q_{m1} - c_m U_m Q_m\|_F \le \|\xi^{(\mathcal{S}_m)} - \xi Q_{m2}\|_F + \frac{\alpha^{(\mathcal{S}_m)}}{d_0}
$$

where $Q_m = Q_{m2}Q_{m1}$.

Step 2.2. (Upper bound for $\|N/\overline{n}_m B_L - B_L^{(\mathcal{S}_m)}\|_{\max}$)

For convenience , denote $h_i = \overline{n}_m B_i^\top \Theta^\top \mathbf{1}_N / N$ and $t_i = \sqrt{\overline{n}_m} f_i / \sqrt{N}$, Then we have

$$
\begin{aligned}
\Big|(\frac{n_m}{N} B_L - B_L^{(\mathcal{S}_m)})_{ij}\Big| &= B_{ij}\Big|\frac{t_i t_j}{\sqrt{h_i h_j}} - \frac{o_i o_j}{\sqrt{a_i a_j}}\Big| \le \Big|\frac{t_i t_j \sqrt{a_i a_j} - o_i o_j \sqrt{h_i h_j}}{\sqrt{h_i h_j a_i a_j}}\Big| \\
&\le \frac{o_i o_j |\sqrt{a_i a_j} - \sqrt{h_i h_j}|}{\sqrt{h_i h_j a_i a_j}} + \frac{|t_i t_j - o_i o_j| \sqrt{a_i a_j}}{\sqrt{h_i h_j a_i a_j}}
\end{aligned}
$$

where recall that $a_i = B_i^\top \Theta^{(\mathcal{S}_m)\top} \mathbf{1}_{\overline{n}_m}$ and $o_i = (\Theta^{(\mathcal{S}_m)\top} \Theta^{(\mathcal{S}_m)})_{ii}^{1/2}$. We then derive the upper bounds for the above two parts respectively. Similar to (B.9), we obtain

$$
\begin{aligned}
|\sqrt{a_i a_j} - \sqrt{h_i h_j}| &\le 2 \max_i \{\sqrt{h}_i, \sqrt{a}_i\} \max_i \sqrt{|a_i - h_i|} + \max_i |a_i - h_i| \\
&\le 3K\overline{n}_m \max\{u_0^{1/2}, u_m^{1/2}\} \alpha^{(\mathcal{S}_m)1/2},
\end{aligned}
$$

where the second inequality is due to that $h_i \le K\overline{n}_m u_0$, $a_i \le K\overline{n}_m u_m$, $|a_i - h_i| \le \overline{n}_m B_i^\top |\Theta^{(\mathcal{S}_m)\top} \mathbf{1}_{\overline{n}_m} / \overline{n}_m - \Theta^\top \mathbf{1}_N / N| \le K\overline{n}_m \alpha^{(\mathcal{S}_m)}$ and $\alpha^{(\mathcal{S}_m)} \le \max\{u_0, u_m\}$. Then we have

$$
\frac{o_i o_j |\sqrt{a_i a_j} - \sqrt{h_i h_j}|}{\sqrt{h_i h_j a_i a_j}} \le \frac{3u_m \max\{u_0^{1/2}, u_m^{1/2}\} \alpha^{(\mathcal{S}_m)1/2}}{K b_{\min}^2 d_0 d_m}. \tag{B.11}
$$

Next, note that

$$
\begin{aligned}
|t_i - o_i| &= \sqrt{\overline{n}_m} \Big| \Big[\big(\frac{\Theta^\top \Theta}{N}\big)^{1/2} - \big(\frac{\Theta^{(\mathcal{S}_m)\top} \Theta^{(\mathcal{S}_m)}}{\overline{n}_m}\big)^{1/2} \Big]_{ii} \Big| \\
&= \sqrt{\overline{n}_m} \frac{|\Theta^\top \Theta / N - \Theta^{(\mathcal{S}_m)\top} \Theta^{(\mathcal{S}_m)} / \overline{n}_m|_{ii}}{(\Theta^\top \Theta / N)_{ii}^{1/2} + (\Theta^{(\mathcal{S}_m)\top} \Theta^{(\mathcal{S}_m)} / \overline{n}_m)_{ii}^{1/2}} \le \frac{\overline{n}_m^{1/2} \alpha^{(\mathcal{S}_m)}}{d_0^{1/2} + d_m^{1/2}}
\end{aligned}
$$

In addition, we have

$$
|t_i t_j - o_i o_j| \le 2 \max_i o_i \max_i |t_i - o_i| + (\max_i |t_i - o_i|)^2 \le \frac{3\overline{n}_m \alpha^{(\mathcal{S}_m)}}{d_0 + d_m}
$$

where the last inequality is due to that $o_i \le \overline{n}_m^{1/2} u_m^{1/2} \le \overline{n}_m^{1/2}$, $\alpha^{(\mathcal{S}_m)} < 1$, and $d_0^{1/2} + d_m^{1/2} \ge d_0 + d_m$ with $d_0, d_m \le 1$, $(d_0^{1/2} + d_m^{1/2})^2 \ge d_0 + d_m$. As a result, the second part is upper bounded by

$$\frac{|t_i t_j - o_i o_j| \sqrt{a_i a_j}}{\sqrt{h_i h_j a_i a_j}} \le \frac{3 u_m \alpha^{(\mathcal{S}_m)}}{K b_{\min}^2 d_0 d_m (d_0 + d_m)}. \tag{B.12}$$

Combining (B.11) and (B.12), we obtain that

$$\left| \left(\frac{n_m}{N} B_L - B_L^{(\mathcal{S}_m)} \right)_{ij} \right| \le \frac{6 \max\{u_0^{1/2}, u_m^{1/2}\} \alpha^{(\mathcal{S}_m)1/2}}{K b_{\min}^2 d_0 d_m (d_0 + d_m)}.$$

where the inequality holds because $\max\{d_0, d_m\} \le 1/2$ for $K \ge 2$ and $\alpha^{(\mathcal{S}_m)} \le \max\{u_0, u_m\}$.

Step 2.3 (Lower bound on τ_m)

Recall that τ_m is the smallest eigenvalue of $B_L^{(\mathcal{S}_m)}$. Similar to the proof of Step 1.3, we could show $\tau_m \ge \sigma_{\min}(B) d_m / (K u_m)$. This completes the proof.

□

Appendix B.4: Proof of Proposition **??**

Proof. Denote $i_k \in \mathcal{C}$ as the original index of the kth pseudo center, $k = 1, ..., K$. Note that for $l + 1 \le i \le \overline{n}_m$ and $k \in \{1,...,K\}$ but $k \ne g_i$

$$\|\widehat{U}_i^{(\mathcal{S}_m)} - \widehat{C}_k^{(\mathcal{S}_m)}\|_2 \ge \|\widehat{C}_{g_i}^{(\mathcal{S}_m)} - \widehat{C}_k^{(\mathcal{S}_m)}\|_2 - \|\widehat{U}_i^{(\mathcal{S}_m)} - \widehat{C}_{g_i}^{(\mathcal{S}_m)}\|_2, \tag{B.13}$$

$$\begin{aligned} \|\widehat{C}_{g_i}^{(\mathcal{S}_m)} - \widehat{C}_k^{(\mathcal{S}_m)}\|_2 \ge{}& \|U_{i_k}^{(\mathcal{S}_m)} - U_{i_{g_i}}^{(\mathcal{S}_m)}\|_2 \\ & - \|Q^{(\mathcal{S}_m)\top} U_{i_k}^{(\mathcal{S}_m)} - \widehat{C}_k^{(\mathcal{S}_m)}\|_2 - \|Q^{(\mathcal{S}_m)\top} U_{i_{g_i}}^{(\mathcal{S}_m)} - \widehat{C}_{g_i}^{(\mathcal{S}_m)}\|_2. \end{aligned} \tag{B.14}$$

According to Statement D.3 in **?** we have

$$\|U_{i_k}^{(\mathcal{S}_m)} - U_{i_{g_i}}^{(\mathcal{S}_m)}\|_2 \geq \sqrt{\frac{2}{D_m}} \tag{B.15}$$

Combining (B.15) and (B.14), we have

$$\|\widehat{C}_{g_i}^{(\mathcal{S}_m)} - \widehat{C}_k^{(\mathcal{S}_m)}\|_2 \geq \sqrt{\frac{2}{D_m}} - 2\zeta_m$$

Further notice that $P_m = \sqrt{2/D_m} - 2\zeta_m$. With the condition (**??**) in Proposition **??**, using (B.13), we have

$$\begin{aligned}\|\widehat{U}_i^{(\mathcal{S}_m)} - \widehat{C}_k^{(\mathcal{S}_m)}\|_2 &\geq \|\widehat{C}_{g_i}^{(\mathcal{S}_m)} - \widehat{C}_k^{(\mathcal{S}_m)}\|_2 - \|\widehat{U}_i^{(\mathcal{S}_m)} - \widehat{C}_{g_i}^{(\mathcal{S}_m)}\|_2 \\ &> P_m - \frac{P_m}{2} = \frac{P_m}{2} > \|\widehat{U}_i^{(\mathcal{S}_m)} - \widehat{C}_{g_i}^{(\mathcal{S}_m)}\|_2,\end{aligned}$$

for any $k \neq g_i$. As a result, node i will be correctly clustered. □

Appendix B.5: Proof of Proposition ??

Proof. The final result holds as long as $\zeta_m = o(\overline{n}_m^{-1/2})$ with probability $1 - \epsilon$. In the following we prove an upper bound on ζ_m first. Before that, we clarify the notations of some matrices that will be used in the following proof.

NOTATIONS: Denote the centers of clustering after implementing k-means on the master server as $\widehat{C} \in \mathbb{R}^{K\times K}$. Recall that $\mathcal{C} = \{i_1, \cdots, i_K\}$ collect indexes of pseudo nodes, where i_k is the index of the node which is closest to the kth center. Correspondingly, let $\widehat{U}_{0c} \overset{\text{def}}{=} (\widehat{U}_{0,i\cdot} : i \in \mathcal{C})^\top \in \mathbb{R}^{K\times K}$ be the mappings of the K pseudo nodes in $\widehat{U}_0$. In addition, let $U_{0c} = (U_{0,i\cdot} : i \in \mathcal{C})^\top \in \mathbb{R}^{K\times K}$, $\widehat{U}_c^{(\mathcal{S}_m)} \overset{\text{def}}{=} (\widehat{U}_{i\cdot}^{(\mathcal{S}_m)}, i \in \mathcal{C})^\top \in \mathbb{R}^{K\times K}$, $U_c^{(\mathcal{S}_m)} \overset{\text{def}}{=} (U_{i\cdot}^{(\mathcal{S}_m)}, i \in \mathcal{C})^\top \in \mathbb{R}^{K\times K}$, and $U_c \overset{\text{def}}{=} (U_{i\cdot}, i \in \mathcal{C})^\top \in \mathbb{R}^{K\times K}$. Let $\widehat{C}_u = (\widehat{C}_{\widehat{g}_i} :$

$1 \leq i \leq l)^\top \in \mathbb{R}^{l \times K}$ collect the k-means centers of clusters where each node belongs to. Next, let $\widehat{\boldsymbol{g}} = (\widehat{g}_1, \cdots, \widehat{g}_l)^\top \in \mathbb{R}^l$.

Note that $\zeta_m \leq \|\widehat{C}^{(\mathcal{S}_m)} - U_c^{(\mathcal{S}_m)}\|_F$. We can bound the distance $\|\widehat{C}^{(\mathcal{S}_m)} - U_c^{(\mathcal{S}_m)}\|_F$ by using the following inequality:

$$\|\widehat{C}^{(\mathcal{S}_m)} - U_c^{(\mathcal{S}_m)} Q^{(\mathcal{S}_m)}\|_F \leq \|\widehat{C}^{(\mathcal{S}_m)} - r_0^{1/2} r_m^{-1/2} U_{0c} Q_m Q^{(\mathcal{S}_m)}\|_F \tag{B.16}$$

$$+ \|r_0^{1/2} r_m^{-1/2} U_{0c} Q_m Q^{(\mathcal{S}_m)} - U_c^{(\mathcal{S}_m)} Q^{(\mathcal{S}_m)}\|_F. \tag{B.17}$$

We now bound (B.16) and (B.17) respectively.

Step 1 Upper bound on (B.16)

First we have

$$\begin{aligned}\|\widehat{C}^{(\mathcal{S}_m)} - r_0^{1/2} r_m^{-1/2} U_{0c} Q_m Q^{(\mathcal{S}_m)}\|_F &= \|r_0^{1/2} r_m^{-1/2} \widehat{U}_{0c} Q_0^\top Q_m - r_0^{1/2} r_m^{-1/2} U_c Q_m\|_F \\ &= r_0^{1/2} r_m^{-1/2} \|\widehat{U}_{0c} Q_0^\top - U_{0c}\|_F = r_0^{1/2} r_m^{-1/2} \|\widehat{U}_{0c} - U_{0c} Q_0\|_F,\end{aligned}$$

where the equality is due to that Q_m, $Q^{(\mathcal{S}_m)}$ are orthogonal matrices. Note that

$$\big\|\widehat{U}_{0c} - U_{0c} Q_0\big\|_F \leq 2\big\|\widehat{U}_{0c} - \widehat{C}\big\|_F + 2\big\|\widehat{C} - U_{0c} Q_0\big\|_F$$

We bound the two right parts in the following three steps.

Step 1.1 Upper bound on $\|\widehat{U}_{0c} - \widehat{C}\|_F$.

Note that rows of $\widehat{U}_{0c}$ are collected as the rows closest to each row in $\widehat{C}$. Then we have

$$\begin{aligned}\|\widehat{U}_{0c} - \widehat{C}\|_F &\leq \frac{1}{\sqrt{d_0 l}} \|\widehat{U}_0 - \widehat{C}_u\|_F \\ &\leq \frac{2}{\sqrt{d_0 l}} (\|\widehat{U}_0 - U_0 Q_0\|_F + \|\widehat{C}_u - U_0 Q_0\|_F)\end{aligned}$$

$\|\widehat{U}_0 - U_0 Q_0\|_F$ has been bounded by Lemma 2 and $\|\widehat{C}_u - U_0 Q_0\|_F$ will be bounded in the next step.

Step 1.2 Bounds of $\|\widehat{C} - U_c Q_0\|_F$ and $\|\widehat{C}_u - U_0 Q_0\|_F$.

First note that $\widehat{C}$ and U_c are K distinct rows extracted from $\widehat{C}_U$ and U_0, respectively. It suffices to obtain the upper bound for each row of $\widehat{C}$, i.e., $\widehat{C}_j$. Denote $\widehat{G}_j \subseteq \{1, ..., l\}$ as the index sets collecting nodes estimated to be in cluster j and G_j as the index sets collecting nodes truly belonging to cluster j, also denote $m_j = |G_j|$ and $\widehat{m}_j = |\widehat{G}_j|$, $j = 1, ..., K$. Further denote $\widehat{U}_0^{(\widehat{G}_j)}$ as a collection of rows indexed in $\widehat{G}_j$ from $\widehat{U}_0$ and $\widehat{U}_{0i}^{(\widehat{G}_j)}$ as the ith row in $\widehat{U}_0^{(\widehat{G}_j)}$. By the definition, we have

$$\begin{aligned}
&\left\|\widehat{C}_j - Q_0^\top C_j\right\|^2 = \left\|\frac{\sum_{i=1}^{\widehat{m}_j} \widehat{U}_{0i}^{(\widehat{G}_j)}}{\widehat{m}_j} - \frac{\sum_{i=1}^{\widehat{m}_j} Q_0^\top U_j}{\widehat{m}_j}\right\|^2 = \frac{1}{\widehat{m}_j^2}\left\|\sum_{i=1}^{\widehat{m}_j}(\widehat{U}_{0i}^{(\widehat{G}_j)} - Q_0^\top U_j)\right\|^2 \\
&\leq \frac{2}{\widehat{m}_j}\sum_{i=1}^{\widehat{m}_j}\|\widehat{U}_{0i}^{(\widehat{G}_j)} - Q_0^\top U_j\|^2 = \frac{2}{\widehat{m}_j}\Big\{\sum_{i\in\widehat{G}_j\cap G_j}\|\widehat{U}_{0i}^{(\widehat{G}_j)} - Q_0^\top U_j\|^2 + \sum_{i\in\widehat{G}_j\setminus G_j}\|\widehat{U}_{0i}^{(\widehat{G}_j)} - Q_0^\top U_j\|^2\Big\}
\end{aligned} \tag{B.18}$$

According to Lemma 3, noting that $\delta_0 \geq b_{\min}$ and $P_0 = lu_0$, an upper bound on the order of $|\widehat{G}_j \setminus G_j|$ and a lower bound on $\widehat{m}_j$, comparing to $|G_j|$, can be obtained as

$$|\widehat{G}_j \setminus G_j| \leq \frac{cu_0 K \log(l/\epsilon_l)}{\lambda_{K,0}^2 b_{\min}}, \quad \widehat{m}_j \geq l d_0 - \frac{cu_0 K \log(l/\epsilon_l)}{\lambda_{K,0}^2 b_{\min}} \overset{\text{def}}{=} \widetilde{m}_j, \tag{B.19}$$

with probability $1 - \epsilon_l$, where c is a finite constant. By the assumption (C2) and (C3) we can derive that $K\log(l/\epsilon_l)/(\lambda_{K,0}^2 b_{\min}) \ll l$. As a result we have $\widetilde{m}_j \geq c_1 l d_0$ asymptotically. Further note that from Lemma 2, $\|\widehat{U}_0 - U_0 Q_0\|_F$ is bounded by

$$\|\widehat{U}_0 - U_0 Q_0\|_F \leq \frac{8\sqrt{6}}{\lambda_{K,0}}\sqrt{\frac{K\log(8l/\epsilon_l)}{l b_{\min}}} \overset{\text{def}}{=} \widetilde{u}_0.$$

since $\delta_0 > lb_{\min}$, where $\lambda_{0,K}$ is the smallest nonzero singular value of $\mathcal{L}_0$. As a result, we have in (B.18) that

$$\sum_{i\in \widehat{G}_j \cap G_j} \|\widehat{U}_{0i}^{(\widehat{G}_j)} - Q_0^\top U_j\|^2 + \sum_{i\in \widehat{G}_j \setminus G_j} \|\widehat{U}_{0i}^{(\widehat{G}_j)} - Q_0^\top U_j\|^2 \le \widetilde{u}_0^2$$

with probability at least $1-\epsilon_l$. Together by using (B.19), we have with probability at least $1-\epsilon_l$

$$\|\widehat{C}_j - Q_0^\top U_j\|^2 = O(\widetilde{u}_0^2/\widetilde{m}_j) = o\big(\frac{J_0 K \log(l/\epsilon_l)}{l^2}\big) \quad \text{with} \quad J_0 = \frac{1}{b_{\min}\lambda_{K,0}^2},$$

for $j = 1, \cdots, K$. Subsequently, the bounds can be obtained as

$$\|\widehat{C} - U_{0c}Q_0\|_F = o\Big(\frac{K J_0^{1/2} \log^{1/2}(l/\epsilon_l)}{l}\Big), \quad \|\widehat{C}_u - U_0 Q_0\|_F = o\Big(\frac{J_0^{1/2} K^{1/2} \log^{1/2}(l/\epsilon_l)}{l^{1/2}}\Big)$$

Step 1.3 Bound of (B.16)

Using the results in Step 1.1 and Step 1.2, considering the number of clusters as a constant and combining the assumptions in Proposition **??**, we have

$$\begin{aligned}
&\|\widehat{C}^{(\mathcal{S}_m)} - r_0^{1/2} r_m^{1/2} U_{0c} Q_m Q^{(\mathcal{S}_m)}\|_F \\
&\le 2 r_0^{1/2} r_m^{-1/2} \big(\frac{1}{\sqrt{d_0 l}} \|\widehat{U}_0 - U_0 Q_0\|_F + \frac{1}{\sqrt{d_0 l}} \|\widehat{C}_u - U_0 Q_0\|_F + \|\widehat{C} - U_{0c} Q_0\|_F\big) \\
&= o\Big(\frac{r_0^{1/2} r_m^{-1/2} K^{1/2} (\log(l/\epsilon_l))^{1/2}}{l d_0^{1/2}} + \frac{r_0^{1/2} r_m^{-1/2} K^{1/2} J_0^{1/2} \log(l/\epsilon_l)}{l d_0^{1/2}} + \frac{r_0^{1/2} r_m^{-1/2} K J_0^{1/2} \log^{1/2}(l/\epsilon_l)}{l}\Big) \\
&= o\Big\{\frac{\log^{1/2}(l/\epsilon_l) K J_0^{1/2}}{l^{1/2} \overline{n}_m^{1/2}}\Big\}
\end{aligned}$$

since $K^2 \log(l/\epsilon_l)/(b_{\min}\lambda_{K,0}^2) \ll l$.

Step 2: Upper bound on (B.17)

According to Proposition **??**, we have

$$\|U^{(\mathcal{S}_m)} - r_m U_m Q_m\|_F \leq \frac{14\sqrt{2}K^2 u_m \max\{u_0^{1/2}, u_m^{1/2}\}\alpha^{(\mathcal{S}_m)1/2}}{\sigma_{\min}(B) b^3 d_0^2 d_m^3 (d_0 + d_m)} + \frac{\alpha^{(\mathcal{S}_m)}}{d_0} \stackrel{\text{def}}{=} \alpha_m$$

Recall that $U^{(\mathcal{S}_m)}$ has K distinct rows, which is recorded in $U_c^{(\mathcal{S}_m)}$. Then it holds that

$$\|r_0^{1/2} r_m^{-1/2} U_{0c} Q_m Q^{(\mathcal{S}_m)} - U_c^{(\mathcal{S}_m)} Q^{(\mathcal{S}_m)}\|_F = \|U_c^{(\mathcal{S}_m)} - r_0^{1/2} r_m^{-1/2} U_{0c} Q_m\|_F$$
$$= \|U_c^{(\mathcal{S}_m)} - r_m^{-1/2} U_c Q_m\|_F \leq \frac{1}{\sqrt{\overline{n}_m d_m}} \|U^{(\mathcal{S}_m)} - r_m U_m Q_m\|_F = o\Big(\frac{\alpha_m}{\overline{n}_m^{1/2} d_m^{1/2}}\Big).$$

Note that $l \gg K$ and $\alpha^{(\mathcal{S}_m)} = o(\sigma_{\min}(B)^2/K^4)$ by Condition (C2) and (C3). By the assumptions, we have and d_0, d_m, u_0, u_m are constants. It leads to that $\xi_m = o(\overline{n}_m^{-1/2})$ a.s., which concludes the proof.

□

APPENDIX C: Proof of Theorems

Appendix C.1: Proof of Theorem **??**

Proof. Define $\widehat{H}^{(\mathcal{S}_m)} = 1/\sqrt{2}(\widehat{U}^{(\mathcal{S}_m)\top}, \widehat{V}^{(\mathcal{S}_m)\top})^\top \in \mathbb{R}^{(n_i+2l)\times l}$. In addition, let $H^{(\mathcal{S}_m)} = 1/\sqrt{2}(U^{(\mathcal{S}_m)\top}, V^{(\mathcal{S}_m)\top})^\top \in \mathbb{R}^{(n_i+2l)\times l}$ be its population version. By Lemma 5.1 of **?** we have

$$\big\|\widehat{H}^{(\mathcal{S}_m)}\widehat{H}^{(\mathcal{S}_m)\top} - H^{(\mathcal{S}_m)} H^{(\mathcal{S}_m)\top}\big\|_F \leq \frac{2K}{\lambda_{K,m}} \big\|\widetilde{L}^{(\mathcal{S}_m)} - \widetilde{\mathcal{L}}^{(\mathcal{S}_m)}\big\|_{\max}. \qquad \text{(C.1)}$$

Note that

$$
\begin{aligned}
&\widehat{H}^{(\mathcal{S}_m)}\widehat{H}^{(\mathcal{S}_m)\top} - H^{(\mathcal{S}_m)}H^{(\mathcal{S}_m)\top} \\
&= \begin{pmatrix} \frac{1}{2}\hat{U}^{(\mathcal{S}_m)}(\hat{U}^{(\mathcal{S}_m)})^\top - \frac{1}{2}U^{(\mathcal{S}_m)}(U^{(\mathcal{S}_m)})^\top & \frac{1}{2}\hat{U}^{(\mathcal{S}_m)}(\hat{V}^{(\mathcal{S}_m)})^\top - \frac{1}{2}U^{(\mathcal{S}_m)}(V^{(\mathcal{S}_m)})^\top \\ \\ \frac{1}{2}\hat{V}^{(\mathcal{S}_m)}\hat{U}^{(\mathcal{S}_m)}\top - \frac{1}{2}V^{(\mathcal{S}_m)}(U^{(\mathcal{S}_m)})^\top & \frac{1}{2}\hat{V}^{(\mathcal{S}_m)}(\hat{V}^{(\mathcal{S}_m)})^\top - \frac{1}{2}V^{(\mathcal{S}_m)}(V^{(\mathcal{S}_m)})^\top \end{pmatrix}
\end{aligned}
$$

This implies $\big\|\widehat{H}^{(\mathcal{S}_m)}\widehat{H}^{(\mathcal{S}_m)\top} - H^{(\mathcal{S}_m)}H^{(\mathcal{S}_m)\top}\big\|_F \geq 1/2\big\|\widehat{U}^{(\mathcal{S}_m)}\widehat{U}^{(\mathcal{S}_m)\top} - U^{(\mathcal{S}_m)}U^{(\mathcal{S}_m)\top}\big\|_F \geq 1/2\big\|\widehat{U}^{(\mathcal{S}_m)} - U^{(\mathcal{S}_m)}Q^{(\mathcal{S}_m)}\big\|_F$. Then the result can be immediately obtained by using (C.1) and Proposition 1. □

Appendix C.2: Proof of Theorem **??**

Proof. Denote E_m as the index sets where nodes are misclustered on server m and let $e^{(\mathcal{S}_m)} = |E_m|$. Using Proposition **??** and Proposition **??**, $e^{(\mathcal{S}_m)}$ can be upper bounded with probability $1-\epsilon_l$ by

$$
e^{(\mathcal{S}_m)} = \sum_{i\in E_m} 1 \leq \frac{4\overline{n}_m}{c^2}\sum_{i\in E_m}\|\widehat{U}_i^{(\mathcal{S}_m)} - \widehat{C}_{g_i}^{(\mathcal{S}_m)}\|_2^2,
$$

where c is a constant. Note that we have we have $\|\widehat{U}_i^{(\mathcal{S}_m)} - \widehat{C}_{g_i}^{(\mathcal{S}_m)}\|_2 \leq \|\widehat{U}_i^{(\mathcal{S}_m)} - Q^{(\mathcal{S}_m)\top}U_i^{(\mathcal{S}_m)}\|_2 + \|Q^{(\mathcal{S}_m)\top}U_i^{(\mathcal{S}_m)} - C_{g_i}^{(\mathcal{S}_m)}\|_2 + \|C_{g_i}^{(\mathcal{S}_m)} - \widehat{C}_{g_i}^{(\mathcal{S}_m)}\|_2$, where $Q^{(\mathcal{S}_m)}$ is defined in Theorem **??**. This yields

$$
\begin{aligned}
e^{(\mathcal{S}_m)} &\leq \frac{12\overline{n}_m}{c^2}\sum_{i\in E_m}\big(\|\widehat{U}_i^{(\mathcal{S}_m)} - Q^{(\mathcal{S}_m)\top}U_i^{(\mathcal{S}_m)}\|_2^2 + \|Q^{(\mathcal{S}_m)\top}U_i^{(\mathcal{S}_m)} - C_{g_i}^{(\mathcal{S}_m)}\|_2^2 \\
&\qquad + \|C_{g_i}^{(\mathcal{S}_m)} - \widehat{C}_{g_i}^{(\mathcal{S}_m)}\|_2^2\big) \\
&\leq \frac{12\overline{n}_m}{c^2}\big(\big\|\widehat{U}^{(\mathcal{S}_m)} - Q^{(\mathcal{S}_m)\top}U^{(\mathcal{S}_m)}\big\|_F^2 + \|U^{(\mathcal{S}_m)} - r_m^{-1/2}U_mQ_m\|_F^2 \\
&\qquad + lu_m\|\widehat{C}^{(\mathcal{S}_m)} - C^{(\mathcal{S}_m)}\|_F^2\big). \qquad \text{(C.2)}
\end{aligned}
$$

Note that $n_m u_m \|\widehat{C}^{(\mathcal{S}_m)} - C^{(\mathcal{S}_m)}\|_F^2 = u_m l \|\widehat{C} - U_{0c} Q_0\|_F^2$ by the proof procedure in Appendix **??**. Further combining the results from Theorem **??**, Proposition **??** and the proof of Proposition **??**, each of which bounds one of the three parts in (C.2), based on the assumptions, we have

$$\mathcal{R}^{(\mathcal{S}_m)} = \frac{e^{(\mathcal{S}_m)}}{\overline{n}_m} \leq o\left(\frac{K^2 \log(l/\epsilon_l)}{b_{\min} l \lambda_{K,0}^2} + \frac{K \log(4(n_m + 2l)/\epsilon_m)}{\lambda_{K,m}\delta_m} + \frac{K^4 \alpha^{(\mathcal{S}_m)}}{\sigma_{\min}(B)^2 b_{\min}^6} \right)$$

with probability at least $1 - \epsilon_m - \epsilon_l$.

$\square$