Supplementary materials for "Decentralized Spectral Graph Analysis: a Flexible Approach with Non-asymptotic Performance Guarantees"

Yu-Cheng Hsiao and I-Hsiang Wang

Graduate Institute of Communication Engineering, National Taiwan University
Email: h411235@gmail.com, ihwang@ntu.edu.tw

1 Problem Formulation

Considering a connected and undirected graph $\mathcal{G}$ with a set of nodes (agents) $V = \{1, 2, ..., N\}$, its structure is captured by a symmetric graph matrix $\boldsymbol{Z} \in \mathbb{R}^{N\times N}$. If the node pair (i, j) does not form an edge in $\mathcal{G}$, then $Z_{ij} = 0$. Otherwise, $Z_{ij} \neq 0$ and $\forall i \in V, Z_{ii}$ could be non-zero.

1.1 Notation

We use $\|\boldsymbol{A}\|_2$ to denote both the Euclidean norm of a vector and the spectral norm of a matrix. $\|\boldsymbol{A}\|_F = (\text{trace}(\boldsymbol{A}^\top \boldsymbol{A}))^{1/2}$ denotes the Frobenius norm of a matrix $\boldsymbol{A}$. For any square matrix $\boldsymbol{A} \in \mathbb{R}^{N\times N}$, $\text{diag}(\boldsymbol{A})$ denotes the matrix obtained by setting all off-diagonal entries of $\boldsymbol{A}$ to 0. For any vector $\boldsymbol{x} \in \mathbb{R}^N$, $\text{diag}(\boldsymbol{x})$ denotes the diagonal matrix constructed by the entries of $\boldsymbol{x}$.

We denote $\mathbb{I}\{\dot{}\}$ as the indicator function. For any two positive integers a, b, we denote $\mathbf{I}_a$ 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 two sequences of real numbers $\{x_N\}$ and $\{y_N\}$, we will write $x_n = o(y_N)$ if $\lim_N x_N/y_N = 0, x_N = \mathcal{O}(y_N)$; if $|x_N/y_N| \leq C$ for all N and some positive C, $x_N = \Omega(y_N)$ if $|x_N/y_N| > C$ for all N and some positive C and $x_N = \Theta(y_N)$ if $x_N = \Omega(y_N)$ and $x_N = \mathcal{O}(y_N)$.

Throughout this work, we use the notation $\lambda_1(\boldsymbol{X}), \lambda_2(\boldsymbol{X}), ..., \lambda_K(\boldsymbol{X})$ to represent the top K eigenvalues of a symmetric matrix $\boldsymbol{X} \in \mathbb{R}^{N\times N}$. Note that eigenvalues are arranged in descending order of magnitude. For simplicity, we refer to the eigenvalues of the graph matrix $\boldsymbol{Z}$ as λ_i for any i. We denote $\{\lambda_i\}_{i=1}^N$ as the eigenvalues of $\boldsymbol{Z}$ and arrange them in the descending order of magnitude such that $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_N|$. The associated eigenvectors are denoted as $\boldsymbol{u}_1, ...\boldsymbol{u}_N \in \mathbb{R}^N$. Let the subscript (i) represent the i_{th} coordinate of vector. Then, scalar $u_{k,(i)}$ represents the i_{th} coordinate in the eigenvector $\boldsymbol{u}_k$ and $u^{(t)}_{k,(i)}$ represents the estimation of scalar $u_{k,(i)}$ at the t_{th} iteration. In addition, the symbol "$\mathtt{H}$" represents the transpose operation.

When it comes to the number of iterations, T_s represents the total number of power iterations for estimating the s_{th} eigenvector and L represents the total number gossip iterations at each **"Gossip"** step. For the weight matrix in **"Gossip"** step, the metropolis weights is employed [1]:

$$\boldsymbol{W}_{ij} \triangleq \begin{cases} 1/\left(1 + \max\{d(i), d(j)\}\right) & j \in \mathtt{N}(i), \\ 1 - \sum_{k\in\mathtt{N}(i)} W_{ik} & i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Here, $\mathtt{N}(i) \triangleq \{j \neq i |\ Z_{ij} \neq 0\}$ as the neighbors of agent i and $d(i)$ as its degree.

Algorithm 1 Decentralized Power Method with Deflation

Input: $\boldsymbol{Z}, \boldsymbol{W}, L, N, K, \{T_i\}_{i=1}^K$

Initialization: $\forall i \in V$, $\{(\boldsymbol{u}_{j,(i)}^{(0)})\}_{j=1}^K \sim \mathcal{N}(\mathbf{0}, \frac{1}{N}\mathbf{I}_K)$

for $r = 1, 2, ..., K$ and $\forall\, i \in V$ **do**

  for $t = 0, 1, 2, .., T_r - 1$ **do**

    for $s = 1, 2, ...r - 1$ **do**

      $\left\{b_{sr,(i)}^{(t+1)}\right\}_{i=1}^N = \mathrm{AC}\left(\left\{u_{s,(i)}^{(T_s)} u_{r,(i)}^{(t)}\right\}_{i=1}^N ; L\right)$ ▷ **Gossip**

    $\hat{u}_{r,(i)}^{(t+1)} = \sum_j Z_{ij} u_{r,(j)}^{(t)} - \sum_{s=1}^{r-1} \lambda_{s,(i)}^{(T_s)} u_{s,(i)}^{(T_s)} N b_{sr,(i)}^{(t+1)}$ ▷ **Aggregation**

    $\left\{c_{(i)}^{(t+1)}\right\}_{i=1}^N = \mathrm{AC}\left(\left\{\left(\hat{u}_{r,(i)}^{(t+1)}\right)^2\right\}_{i=1}^N ; L\right)$ ▷ **Gossip**

    $\lambda_{r,(i)}^{(t+1)} = \sqrt{N\, c_{(i)}^{(t+1)}}$

    $u_{r,(i)}^{(t+1)} = \hat{u}_{r,(i)}^{(t+1)} / \sqrt{N\, c_{(i)}^{(t+1)}}$ ▷ **Local normalization**

1.2 Assumption

We make the following assumption in the paper:

Assumption 1. *The top $(K+1)$ eigenvalues of symmetric matrix $\boldsymbol{Z}$ are unique. Moreover, the top K eigenvalues of symmetric matrix $\boldsymbol{Z}$ are positive.*

The assumption of distinct eigenvalues is crucial for the deflation process, as it eliminates the effect of a specific eigenpair. Fortunately, all eigenvalues of graph matrices are unique with high probability if the underlying graph is from several random graph models (**Lemma 1.1**). In addition, the power method only estimates eigenvalue magnitudes, but it's noteworthy that many graph matrices have non-negative leading eigenvalues, including the normalized Laplacian matrix for undirected graphs and the adjacency matrix for graphs with assortative communities. For example, the adjacency matrix for graphs from Stochastic Block Model with K communities (**Lemma 3.6**).

Lemma 1.1. *([2, Theorem 5.1]) For any fixed $r, \nu \geq 0$ and sufficiently large N the following holds. Let $\xi_{ij}, 1 \leq i < j \leq N$ be independent (complex or real) random variables such that $\mathbb{P}(\xi_{ij} = x) \leq 1 - \nu$ for any $1 \leq i < j \leq N$ and $x \in \mathbb{R}$. Let $\xi_{ii}, 1 \leq i \leq N$ be real random variables that are independent of the $\xi_{ij}, 1 \leq i < j \leq N$. Set $\xi_{ji} = \overline{\xi_{ij}}$ for $1 \leq i < j \leq N$. Then the spectrum of the matrix $(\xi_{ij})_{1 \leq i < j \leq N}$ is simple with probability at least $1 - N^{-r}$.*

Remark 1.1. *In Lemma 1.1, $\overline{\xi_{ij}}$ denotes the conjugate of ξ_{ij}. Also, the spectrum is simple, which means all eigenvalues are distinct.*

2 Non-asymptotic guarantees

To derive the estimation error for individual eigenvector, we first define the tolerance for estimation error of i_{th} eigenvector as ϵ_i. Given a finite number of $\{T_i\}_{i=1}^K$ and L, our goal is to show the followings hold :

$$\forall i \in [K], \|\boldsymbol{u}_i - \boldsymbol{u}_i^{(T_i)}\|_2 \leq \epsilon_i.$$

Here, $\epsilon_i < \frac{1}{2}$ for all i. T_i represents the number of power iterations for estimating the i_{th} eigenvector of symmetric matrix $\boldsymbol{Z}$ and L denotes the number of gossip iterations at each "**Gossip**" step. Moreover $\boldsymbol{u}_i^{(T_i)}$ for all i denotes the outputs of **Algorithm 1**.

2.1 K=2

We begin with the case $K=2$. For the centralized power method, at each iteration $t=0,1,...T_1-1$, the algorithm can be written as follows:

$$\begin{aligned}
\hat{\boldsymbol{q}}_1^{(t+1)} &= \boldsymbol{Z}\boldsymbol{q}_1^{(t)} \\
\boldsymbol{q}_1^{(t+1)} &= \hat{\boldsymbol{q}}_1^{(t+1)}/\|\hat{\boldsymbol{q}}_1^{(t+1)}\|_2 \\
\gamma_1^{(t+1)} &= \|\hat{\boldsymbol{q}}_1^{(t+1)}\|_2
\end{aligned} \tag{2}$$

To estimate $(\lambda_2, \boldsymbol{u}_2)$:

$$\begin{aligned}
\hat{\boldsymbol{q}}_2^{(t+1)} &= \left(\boldsymbol{Z} - \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^{\mathtt{H}}\right) \boldsymbol{q}_2^{(t)} \\
\boldsymbol{q}_2^{(t+1)} &= \hat{\boldsymbol{q}}_2^{(t+1)}/\|\hat{\boldsymbol{q}}_2^{(t+1)}\|_2 \\
\gamma_2^{(t+1)} &= \|\hat{\boldsymbol{q}}_2^{(t+1)}\|_2
\end{aligned} \tag{3}$$

Here, $\gamma_i^{(T_i)}$ and $\boldsymbol{q}_i^{(T_i)}$ represent the estimation of i_{th} eigenvalue and the resulting vector by centralized power method, respectively. Turing to the decentralized counterpart, **Algorithm 1** with $K=2$ can be expressed as:

$$\begin{aligned}
\hat{\boldsymbol{u}}_2^{(t+1)} &= \left(\boldsymbol{Z} - \boldsymbol{\Lambda}_1^{(T_1)} \boldsymbol{u}_1^{(T_1)} (\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}}\right) \boldsymbol{u}_2^{(t)} \\
&= \boldsymbol{Z}\boldsymbol{u}_2^{(t)} - \boldsymbol{\Lambda}_1^{(T_1)} \boldsymbol{u}_1^{(T_1)} \underbrace{(\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}} \boldsymbol{u}_2^{(t)}}_{\text{Estimated by \textbf{``Gossip''} step}}
\end{aligned} \tag{4}$$

where the diagonal matrix $\boldsymbol{\Lambda}_1^{(T_1)} = [\lambda_{1,(1)}^{(T_1)}, ..., \lambda_{1,(N)}^{(T_1)}]\mathbf{I}_N$ and $\lambda_{1,(i)}^{(T_1)}$ represents the estimated eigenvalue at node i. Moreover, we can rewrite Equation (4) as follows:

$$\begin{aligned}
\hat{\boldsymbol{u}}_2^{(t+1)} &= \boldsymbol{Z}\boldsymbol{u}_2^{(t)} - \boldsymbol{\Lambda}_1^{(T_1)} \boldsymbol{u}_1^{(T_1)} (\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}} \boldsymbol{u}_2^{(t)} \\
&= \boldsymbol{Z}\left(\boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)} + \tilde{\boldsymbol{u}}_2^{(t)}\right) - \gamma_1^{(T_1)} \boldsymbol{q}_1^{(T_1)} (\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}} \boldsymbol{u}_2^{(t)} + \underbrace{\gamma_1^{(T_1)} \boldsymbol{q}_1^{(T_1)} (\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}} \boldsymbol{u}_2^{(t)} - \boldsymbol{\Lambda}_1^{(T_1)} \boldsymbol{u}_1^{(T_1)} (\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}} \boldsymbol{u}_2^{(t)}}_{\boldsymbol{G}_d^{(t)}} \\
&= \underbrace{\left(\boldsymbol{Z} - \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^{\mathtt{H}}\right)}_{\text{ideal deflated matrix}} \tilde{\boldsymbol{u}}_2^{(t)} + \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^{\mathtt{H}} \tilde{\boldsymbol{u}}_2^{(t)} + \boldsymbol{Z}\left(\boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)}\right) - \gamma_1^{(T_1)} \boldsymbol{q}_1^{(T_1)} (\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}} \boldsymbol{u}_2^{(t)} + \boldsymbol{G}_d^{(t)} \\
&= \boldsymbol{\mathcal{D}}^{(t)} \tilde{\boldsymbol{u}}_2^{(t)} + \underbrace{\left(\boldsymbol{Z} - \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^{\mathtt{H}}\right)\left(\boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)}\right)}_{\boldsymbol{G}_{cd}^{(t)}} + \underbrace{\lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^{\mathtt{H}} \boldsymbol{u}_2^{(t)} - \gamma_1^{(T_1)} \boldsymbol{q}_1^{(T_1)} (\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}} \boldsymbol{u}_2^{(t)}}_{\boldsymbol{G}_c^{(t)}} + \boldsymbol{G}_d^{(t)} \\
&= \boldsymbol{\mathcal{D}}^{(t)} \tilde{\boldsymbol{u}}_2^{(t)} + \boldsymbol{G}_{cd}^{(t)} + \boldsymbol{G}_c^{(t)} + \boldsymbol{G}_d^{(t)} \\
&= \boldsymbol{\mathcal{D}}^{(t)} \tilde{\boldsymbol{u}}_2^{(t)} + \boldsymbol{G}^{(t)}
\end{aligned} \tag{5}$$

In Equation (5), $\tilde{\boldsymbol{u}}_2^{(t)} = \dfrac{\hat{\boldsymbol{u}}_1^{(t)}}{\|\hat{\boldsymbol{u}}_1^{(t)}\|_2}$ denotes the ideal normalized vector and the decentralized estimation $\boldsymbol{u}_1^{(T_1)}$ is used to construct the deflated matrix $\boldsymbol{Z} - \lambda_1^{(T_1)} \boldsymbol{u}_1^{(T_1)} (\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}}$ for estimating $(\lambda_2, \boldsymbol{u}_2)$, that is, the decentralization is equivalent to adding a perturbation $\boldsymbol{G}^{(t)}$ to the ideal updates $\boldsymbol{\mathcal{D}}^{(t)} \tilde{\boldsymbol{u}}_2^{(t)} \triangleq \left(\boldsymbol{Z} - \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^{\mathtt{H}}\right) \tilde{\boldsymbol{u}}_2^{(t)}$ at each iteration t. Fortunately, when the perturbation $\boldsymbol{G}^{(t)}$ is less than a specified threshold, the power method with deflation process still converges, as described in **Corollary 1**.

Remark 2.1. *Since Assumption 1 assumes the top K eigenvalues are positive, the sign of λ_i is the same as the sign of eigenvalue $\gamma_i^{(t)}$ and $\lambda_i^{(t)}$ for all t and all $i = 1, 2, ..., K$.*

Theorem 1. *([3, Theorem 2.4]) Let $\boldsymbol{u} \in \mathbb{R}^N$ represent the top singular vector of $\boldsymbol{B} \in \mathbb{R}^{N\times N}$. Suppose $\boldsymbol{x}^{(0)}$ is an random vector. Furthermore, suppose at every step of power method with ideal normalization we have*

$$5\left\|\boldsymbol{G}^{(t)}\right\|_2 \leq \rho\left(|\lambda_1(\boldsymbol{B})| - |\lambda_2(\boldsymbol{B})|\right) \text{ and } 5\left\|\boldsymbol{u}^{\mathtt{H}}\boldsymbol{G}^{(t)}\right\|_2 \leq \left(|\lambda_k(\boldsymbol{B})| - |\lambda_{k+1}(\boldsymbol{B})|\right)\cos\theta^{(0)},$$

where $\cos\theta^{(0)} = \boldsymbol{u}^{\mathtt{H}}\boldsymbol{x}^{(0)}$ *and the perturbation matrix* $\boldsymbol{G}^{(t)}$*. Then there exists an*

$$T_2 = \Omega\left(\frac{|\lambda_1(\boldsymbol{B})|}{|\lambda_1(\boldsymbol{B})| - |\lambda_2(\boldsymbol{B})|}\log\frac{\tan\theta^{(0)}}{\rho}\right)$$

so that after T_2 *steps we have*

$$\left\|\left(\boldsymbol{I} - \boldsymbol{X}^{(T_2)}(\boldsymbol{X}^{(T_2)})^{\mathtt{H}}\right)\boldsymbol{u}\right\|_2 = \sqrt{1 - \boldsymbol{u}^{\mathtt{H}}\boldsymbol{X}^{(T_2)}(\boldsymbol{X}^{(T_2)})^{\mathtt{H}}\boldsymbol{u}} \leq \rho.$$

Corollary 1. *For* $\cos\theta_2^{(0)} = \boldsymbol{u}_2^{\mathtt{H}}\boldsymbol{u}_2^{(0)} \geq 0$ *and* $\rho < \frac{1}{2}$*, executing the power method based on Equation (5), we have*

$$\|\boldsymbol{u}_2 - \tilde{\boldsymbol{u}}_2^{(T_2)}\|_2 \leq \sqrt{2}\rho,\ T_2 = \Omega\left(\frac{\lambda_2}{\lambda_2 - \lambda_3}\log\left(\frac{\tan\theta_2^{(0)}}{\rho}\right)\right),$$

as long as

$$5\|\boldsymbol{G}^{(t)}\|_2 \leq \min\{\rho, \cos\theta_2^{(0)}\}\left(\lambda_2 - \lambda_3\right), \tag{6}$$

where $\boldsymbol{G}$ *represents the perturbation matrix at each iteration* t*.*

Proof. We directly apply the [3, Theorem 2.4], which is presented in **Theorem 1**. In the context of Equation 5, the initial vector $\boldsymbol{x}_0 = \boldsymbol{u}_2^{(0)}$ and the ideal eigenvector vector $\boldsymbol{u} = \boldsymbol{u}_2$. Also, the deflated matrix is

$$\boldsymbol{B} = \boldsymbol{Z} - \lambda_1\boldsymbol{u}_1\boldsymbol{u}_1^{\mathtt{H}},$$

and the perturbation matrix is

$$\boldsymbol{G}^{(t)} = \left(\boldsymbol{Z} - \lambda_1\boldsymbol{u}_1\boldsymbol{u}_1^{\mathtt{H}}\right)\left(\boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)}\right) + \boldsymbol{G}_c^{(t)} - \boldsymbol{G}_d^{(t)}.$$

Since $\lambda_1(\boldsymbol{B}) = \lambda_2(\boldsymbol{Z})$ and $\lambda_2(\boldsymbol{B}) = \lambda_3(\boldsymbol{Z})$, we can derive the sufficient condition as follows:

$$5\left\|\boldsymbol{G}^{(t)}\right\|_2 \leq \rho\left(|\lambda_1(\boldsymbol{B})| - |\lambda_2(\boldsymbol{B})|\right) = 5\left\|\boldsymbol{G}^{(t)}\right\|_2 \leq \rho\left(|\lambda_2 - |\lambda_3|\right)$$

and

$$5\left\|\boldsymbol{u}^{\mathtt{H}}\boldsymbol{G}^{(t)}\right\|_2 \leq 5\left\|\boldsymbol{u}^{\mathtt{H}}\right\|_2\left\|\boldsymbol{G}^{(t)}\right\|_2 = 5\left\|\boldsymbol{G}^{(t)}\right\|_2 \leq (|\lambda_2| - |\lambda_3|)\cos\theta_2^{(0)},$$

where $\cos\theta_2^{(0)} = \boldsymbol{u}_2^{\mathtt{H}}\boldsymbol{u}_2^{(0)}$. Therefore, we can derive the sufficient condition as:

$$5\left\|\boldsymbol{G}^{(t)}\right\|_2 \leq (|\lambda_2| - |\lambda_3|)\min\{\rho, \cos\theta_2^{(0)}\}.$$

By the definition, $\tilde{\boldsymbol{u}}_2^{(t)}$ and $\boldsymbol{u}_2$ are unit vectors, so we have:

$$\begin{aligned}
&\|\tilde{\boldsymbol{u}}_2^{(t)} - \boldsymbol{u}_2\|_2^2 = 2\left(1 - \boldsymbol{u}_2^{\mathtt{H}}\tilde{\boldsymbol{u}}_2^{(t)}\right) \leq 2\left(1 - \left(\boldsymbol{u}_2^{\mathtt{H}}\tilde{\boldsymbol{u}}_2^{(t)}\right)^2\right) = 2\rho^2 \\
&\|\tilde{\boldsymbol{u}}_2^{(t)} - \boldsymbol{u}_2\|_2 \leq \sqrt{2}\rho
\end{aligned}$$

as long as

$$T_2 = \Omega\left(\frac{\lambda_2}{\lambda_2 - \lambda_3}\log\left(\frac{\tan\theta_2^{(0)}}{\rho}\right)\right).$$

□

Now, our goal is to show Equation (6) is satisfied with a finite number of T_1, T_2 and L. According to Equation (5), $\boldsymbol{G}^{(t)}$ can be decomposed as follows:

$$\begin{aligned}
\|\boldsymbol{G}^{(t)}\|_2 &= \|_2 \left(\boldsymbol{Z} - \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^{\mathtt{H}}\right) \left(\boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)}\right) + \underbrace{\lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^{\mathtt{H}} \boldsymbol{u}_2^{(t)} - \gamma_1^{(T_1)} \boldsymbol{q}_1^{(T_1)} (\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}} \boldsymbol{u}_2^{(t)}}_{\boldsymbol{G}_c^{(t)}} + \boldsymbol{G}_d^{(t)} \|_2 \\
&\leq \left\| \left(\boldsymbol{Z} - \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^{\mathtt{H}}\right) \left(\boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)}\right) \right\|_2 + \left\| \boldsymbol{G}_c^{(t)} \right\|_2 + \left\| \boldsymbol{G}_d^{(t)} \right\|_2 \\
&\leq \left\| \boldsymbol{Z} - \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^{\mathtt{H}} \right\|_2 \left\| \boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)} \right\|_2 + \left\| \boldsymbol{G}_c^{(t)} \right\|_2 + \left\| \boldsymbol{G}_d^{(t)} \right\|_2 \\
&\leq \lambda_2 \left\| \boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)} \right\|_2 + \left\| \boldsymbol{G}_c^{(t)} \right\|_2 + \left\| \boldsymbol{G}_d^{(t)} \right\|_2,
\end{aligned} \tag{7}$$

where the last inequality holds since the largest eigenvalue of ideal deflated matrix $(\boldsymbol{Z} - \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^{\mathtt{H}})$ is λ_2. Moreover, we will demonstrate the followings in **lemma 2.12**:

$$\begin{aligned}
\left\| \boldsymbol{G}_{cd}^{(T_2)} \right\|_2 &\leq |\lambda_2| \left\| \boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)} \right\|_2 \leq C_2 \lambda_1^2 |\lambda_2| \phi^L \\
\|\boldsymbol{G}_c^{(t)}\|_2 &= \mathcal{O}\left(\left(\tan \theta_1^{(0)}\right) |\lambda_1| \left| \frac{\lambda_2}{\lambda_1} \right|^{T_1} \right), \\
\|\boldsymbol{G}_d^{(t)}\|_2 &= \mathcal{O}\left(N |\lambda_1^3| \phi^L \left| \frac{\lambda_1}{\lambda_N} \right|^{T_1 - 1} \right).
\end{aligned}$$

where ϕ represents the second largest eigenvalue of weight matrix $\boldsymbol{W}$. To satisfy Equation (6), the sufficient number of gossip iterations is

$$L = \Omega\left(\frac{1}{\log(\phi^{-1})} \frac{\log\left(\left|\frac{\lambda_1}{\lambda_N}\right|\right)}{\log\left(\frac{\lambda_1}{\lambda_2}\right)} \log\left(\frac{\left(\tan \theta_1^{(0)}\right) \sqrt{N} |\lambda_1^3|}{(\lambda_2 - \lambda_3) \min\{\rho, \frac{1}{\tau\sqrt{N}}\}} \right) \right)$$

and the number of power iteration are

$$T_1 = \Omega\left(\frac{1}{\log(\lambda_1) - \log(\lambda_2)} \log\left(\frac{\left(\tan \theta_1^{(0)}\right) |\lambda_1|}{(\lambda_2 - \lambda_3) \min\{\rho, \frac{1}{\tau\sqrt{N}}\}} \right) \right), \; T_2 = \Omega\left(\frac{\lambda_2}{\lambda_2 - \lambda_3} \log\left(\frac{\tan \theta_2^{(0)}}{\rho} \right) \right).$$

Then, apply **Lemma 2.1** to $\tan\theta_1^{(0)}, \tan\theta_2^{(0)}, \cos\theta_2^{(0)}$. The formal convergence arguments are summarized in **Theorem 2**.

Lemma 2.1. *([3, Lemma 2.5]) For $\boldsymbol{U}^{(0)} = [\mathbf{u}_1^{(0)}, ..., \boldsymbol{u}_K^{(0)}] \in \mathbb{R}^{N \times K}$ with independent and identically distributed Gaussian entries based on $\mathcal{N}(0, \frac{1}{N})$ and $\tau > 1$, with $1 - \tau^{-\Omega(1)} - e^{-\Omega(N)}$ probability, we have*

$$\begin{aligned}
\forall i = 1, 2, ..., K, \; \tan \theta_i^{(0)} &\leq \frac{\tau\sqrt{N}}{\sqrt{K} - \sqrt{K-1}} \leq 2\tau\sqrt{NK} \\
\cos \theta_i^{(0)} &\geq \frac{\sqrt{K} - \sqrt{K-1}}{\tau\sqrt{N}} \geq \frac{1}{2\tau\sqrt{NK}}
\end{aligned}$$

where $\theta_i^{(0)} = \arccos\left(\frac{\boldsymbol{u}_i^{\mathtt{H}} \boldsymbol{u}_i^{(0)}}{\|\boldsymbol{u}_i^{(0)}\|_2}\right)$ and $\boldsymbol{u}_i$ denotes the i_{th} eigenvector of symmetric matrix $\boldsymbol{Z}$. Note that

$$\sqrt{K} - \sqrt{K-1} \geq \frac{1}{2\sqrt{K}}.$$

Proof. We denote the top K eigenvectors of matrix $\boldsymbol{Z}$ as $\boldsymbol{U} \in \mathbb{R}^{N \times K}$, which has orthonormal columns $\boldsymbol{u}_1, ..., \boldsymbol{u}_K$. Also, $\boldsymbol{\mathcal{U}}^{(0)} = [\mathtt{u}_1^{(0)}, ..., \mathtt{u}_K^{(0)}] \in \mathbb{R}^{N \times K}$ with independent and identically distributed Gaussian entries based on $\mathcal{N}(0, 1)$.

Let us define $\theta_i^{(0)} = \arccos\left(\frac{\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{u}_i^{(0)}}{\|\boldsymbol{u}_i^{(0)}\|_2}\right)$ and the null space $\boldsymbol{V}_i = (\boldsymbol{u}_i)^{\perp} \in \mathbb{R}^{N\times(N-1)}$, then we have

$$\forall i = 1, ..., K,\ \tan\theta_i^{(0)} = \frac{\|\boldsymbol{V}_i^{\mathrm{H}}\boldsymbol{u}_i^{(0)}\|_2}{\|\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{u}_i^{(0)}\|_2} \leq \frac{\|\boldsymbol{u}_i^{(0)}\|_2}{\|\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{u}_i^{(0)}\|_2} = \frac{\|\sqrt{N}\boldsymbol{u}_i^{(0)}\|_2}{\|\sqrt{N}\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{u}_i^{(0)}\|_2} \leq \frac{\sqrt{N}\|\boldsymbol{u}_i^{(0)}\|_2}{\sigma_K\left(\boldsymbol{U}^{\mathrm{H}}\boldsymbol{\mathcal{U}}^{(0)}\right)}.$$

Here, we use the scaling property of variance to apply [4, Theorem 1.1], which provides a lower bound for the smallest singular value of matrix with independent and identically distributed Gaussian entries based on $\mathcal{N}(0,1)$. Since the smallest singular value $\sigma_K\left(\boldsymbol{U}^{\mathrm{H}}\boldsymbol{\mathcal{U}}^{(0)}\right)$ can be expressed as

$$\sigma_K\left(\boldsymbol{U}^{\mathrm{H}}\boldsymbol{\mathcal{U}}^{(0)}\right) = \inf_{\boldsymbol{x}\in\mathbb{R}^N:\|\boldsymbol{x}\|_2=1}\left\|\boldsymbol{U}^{\mathrm{H}}\boldsymbol{\mathcal{U}}^{(0)}\boldsymbol{x}\right\|_2,$$

we can derive $\|\sqrt{N}\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{u}_i^{(0)}\|_2 \geq \sigma_K\left(\boldsymbol{U}^{\mathrm{H}}\boldsymbol{\mathcal{U}}^{(0)}\right)$. Furthermore, using [4, Theorem 1.1], with probability at least $1 - \tau^{-\Omega(1)} - e^{-\Omega(N)}$ for some $\tau > 1$, we have

$$\sigma_K\left(\boldsymbol{U}^{\mathrm{H}}\boldsymbol{\mathcal{U}}^{(0)}\right) \geq \frac{1}{\tau}\left(\sqrt{K} - \sqrt{K-1}\right).$$

On the other hand, $\forall i, \|\boldsymbol{u}_i^{(0)}\|_2 \geq \frac{3}{2}$ with probability $e^{-\Omega(N)}$, which is shown in **Lemma 4.1**. Finally, we have

$$\forall i = 1, ..., K,\ \tan\theta_i^{(0)} \leq \frac{\sqrt{N}\|\boldsymbol{u}_i^{(0)}\|_2}{\sigma_K\left(\boldsymbol{U}^{\mathrm{H}}\boldsymbol{\mathcal{U}}^{(0)}\right)} \leq \frac{\tau\frac{3}{2}\sqrt{N}}{\sqrt{K} - \sqrt{K-1}}$$

with probability $1 - \tau^{-\Omega(1)} - e^{-\Omega(N)}$. Then, rescale τ to obtain the desired result. Meanwhile, we also have

$$\forall i = 1, ..., K,\ \cos\theta_i^{(0)} = \frac{\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{u}_i^{(0)}}{\|\boldsymbol{u}_i^{(0)}\|_2} \geq \frac{\sigma_K\left(\boldsymbol{U}^{\mathrm{H}}\boldsymbol{\mathcal{U}}^{(0)}\right)}{\sqrt{N}\|\boldsymbol{u}_i^{(0)}\|_2} \geq \frac{\sqrt{K} - \sqrt{K-1}}{\tau\sqrt{N}}.$$

□

Theorem 2. *For Algorithm 1 with $K = 2, \tau > 1$ and $\rho < \frac{1}{2}$, with $1 - \tau^{-\Omega(1)} - e^{-\Omega(N)}$ probability, we have*

$$\|\boldsymbol{u}_1 - \boldsymbol{u}_1^{(T_1)}\|_2 = o(\rho),\ \|\boldsymbol{u}_2 - \boldsymbol{u}_2^{(T_2)}\|_2 \leq \sqrt{2}\rho + o(\rho),$$

when the number of power iterations

$$T_1 = \Omega\left(\frac{1}{\log(\lambda_1) - \log(\lambda_2)}\log\left(\frac{\tau\sqrt{N}|\lambda_1|}{(\lambda_2 - \lambda_3)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}}\right)\right),\ T_2 = \Omega\left(\frac{\lambda_2}{\lambda_2 - \lambda_3}\log\left(\frac{\tau\sqrt{N}}{\rho}\right)\right).$$

and the number of gossip iterations

$$L = \Omega\left(\frac{1}{\log(\phi^{-1})}\frac{\log\left(\left|\frac{\lambda_1}{\lambda_N}\right|\right)}{\log\left(\frac{\lambda_1}{\lambda_2}\right)}\log\left(\frac{\tau N|\lambda_1^3|}{(\lambda_2 - \lambda_3)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}}\right)\right),$$

for $\tau > 1$. Here, ϕ represents the second largest eigenvalue of $\boldsymbol{W}$. Note that the probability comes from the random initialization.

Sketch of proof. From **Lemma 2.12**, we can obtain $\|\boldsymbol{G}^{(t)}\|_2 = \mathcal{O}\left((\lambda_2 - \lambda_3)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}\right)$ when the number of power iterations

$$T_1 = \Omega\left(\frac{1}{\log(\lambda_1) - \log(\lambda_2)}\log\left(\frac{\tau\sqrt{N}|\lambda_1|}{(\lambda_2 - \lambda_3)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}}\right)\right),\ T_2 = \Omega\left(\frac{\lambda_2}{\lambda_2 - \lambda_3}\log\left(\frac{\tau\sqrt{N}}{\rho}\right)\right).$$

and for $\tau > 1$, the number of gossip iterations

$$L = \Omega\left(\frac{1}{\log\left(\phi^{-1}\right)}\frac{\log\left(\left|\frac{\lambda_1}{\lambda_N}\right|\right)}{\log\left(\frac{\lambda_1}{\lambda_2}\right)}\log\left(\frac{\tau N|\lambda_1^3|}{(\lambda_2-\lambda_3)\min\{\rho,\frac{1}{\tau\sqrt{N}}\}}\right)\right).$$

With above T_1, T_2 and L, we can derive the following results:

$$\begin{aligned}
\|\boldsymbol{u}_1-\boldsymbol{u}_1^{(T_1)}\|_2 &\leq \|\boldsymbol{u}_1-\boldsymbol{q}_1^{(T_1)}\|_2+\|\boldsymbol{q}_1^{(T_1)}-\boldsymbol{u}_1^{(T_1)}\|_2\\
&\leq \underbrace{\sqrt{2}\tan\left(\theta_0\right)\left|\frac{\lambda_2}{\lambda_1}\right|^{T_1}}_{\textbf{Lemma 4.3}}+\underbrace{\mathcal{O}\left(\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}\lambda_1^2\phi^L\right)}_{\textbf{Lemma 2.5}}\\
&\leq \underbrace{\sqrt{2}\tau\sqrt{N}\left|\frac{\lambda_2}{\lambda_1}\right|^{T_1}}_{\textbf{Lemma 2.1}}+\mathcal{O}\left(\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}\lambda_1^2\phi^L\right)\underset{\textbf{Lemma 2.11}}{=}\mathcal{O}\left(\frac{\min\{\rho,\frac{1}{\tau\sqrt{N}}\}\}}{N}\right)=o\left(\rho\right).
\end{aligned}\tag{8}$$

Now, we aim to bound $\|\boldsymbol{u}_2^{(T_2)}-\boldsymbol{u}_2\|_2$, which can be decomposed as two components as:

$$\|\boldsymbol{u}_2^{(T_2)}-\boldsymbol{u}_2\|_2=\|\boldsymbol{u}_2-\tilde{\boldsymbol{u}}_2^{(T_2)}\|_2+\|\tilde{\boldsymbol{u}}_2^{(T_2)}-\boldsymbol{u}_2^{(T_2)}\|_2\tag{9}$$

Here, $\tilde{\boldsymbol{u}}_2^{(T_2)}=\frac{\hat{\boldsymbol{u}}_2^{(T_2)}}{\|\hat{\boldsymbol{u}}_2^{(T_2)}\|_2}$ denotes the ideal normalized vector. For the first term in Equation (9), **Theorem 1** is used to bound the angle between $\boldsymbol{u}_2$ and $\boldsymbol{q}_2^{(T_2)}$, that is, $\sqrt{1-\left((\tilde{\boldsymbol{u}}_2^{(T_2)})^{\mathtt{H}}\boldsymbol{u}_2\right)^2}\leq\rho$. Here, we assume $\boldsymbol{u}_2^{\mathtt{H}}\tilde{\boldsymbol{u}}_2^{(T_2)}\geq 0$. Then, we have

$$\begin{aligned}
\|\boldsymbol{u}_2-\tilde{\boldsymbol{u}}_2^{(T_2)}\|_2 &= \sqrt{\|\boldsymbol{u}_2\|_2^2+\|\tilde{\boldsymbol{u}}_2^{(T_2)}\|_2^2-2\boldsymbol{u}_2^{\mathtt{H}}\tilde{\boldsymbol{u}}_2^{(T_2)}}\\
&= \sqrt{2\left(1-\boldsymbol{u}_2^{\mathtt{H}}\tilde{\boldsymbol{u}}_2^{(T_2)}\right)}\\
&\leq \sqrt{2\left(1-\left(\boldsymbol{u}_2^{\mathtt{H}}\tilde{\boldsymbol{u}}_2^{(T_2)}\right)^2\right)}\\
&\leq \sqrt{2}\rho,
\end{aligned}\tag{10}$$

where the first inequality holds due to $\|\boldsymbol{u}_2^{\mathtt{H}}\tilde{\boldsymbol{u}}_2^{(T_2)}\|_2\leq\|\boldsymbol{u}_2^{\mathtt{H}}\|_2\|\tilde{\boldsymbol{u}}_2^{(T_2)}\|_2\leq 1$ Now, we tackle the second term in Equation (9). From **Lemma 2.10**, we have

$$\|\boldsymbol{u}_2^{(T_2)}-\boldsymbol{q}_2^{(T_2)}\|_2=\mathcal{O}\left(\lambda_1^2\phi^L\right).$$

Substituting T_1, T_2 and L in conjunction with $\tan\theta_1^{(0)}\leq\tau\sqrt{N}$, we have

$$\|\boldsymbol{u}_2^{(T_2)}-\tilde{\boldsymbol{u}}_2^{(T_2)}\|_2=\mathcal{O}\left(\frac{\lambda_2-\lambda_3}{N\lambda_1}\min\{\rho,\frac{1}{\tau\sqrt{N}\}}\right).$$

Finally, we can derive the following upper bound.

$$\begin{aligned}
\|\boldsymbol{u}_2^{(T_2)}-\boldsymbol{u}_2\|_2 &\leq \|\boldsymbol{u}_2-\tilde{\boldsymbol{u}}_2^{(T_2)}\|_2+\|\tilde{\boldsymbol{u}}_2^{(T_2)}-\boldsymbol{u}_2^{(T_2)}\|_2\\
&\leq \sqrt{2}\rho+\mathcal{O}\left(\frac{\lambda_2-\lambda_3}{N\lambda_1}\min\{\rho,\frac{1}{\tau\sqrt{N}\}}\right)\\
&= \sqrt{2}\rho+o\left(\rho\right).
\end{aligned}\tag{11}$$

In addition, **Lemma 2.5** shows the estimated eigenvalue is close to the eigenvalue as

$$\max_{i=1,\ldots,N}\left|\lambda_1-\lambda_{1,(i)}^{(T_1)}\right|=\mathcal{O}\left(\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}\sqrt{N}|\lambda_N|\lambda_1^2\phi^L\right).$$

By **Lemma 2.11** and **Lemma 2.12**, we have

$$\|\boldsymbol{G}_d^{(t)}\|_2 = \mathcal{O}\left(\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1} N|\lambda_1^3|\phi^L\right) = \mathcal{O}\left((\lambda_2-\lambda_3)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}\right),$$

which implies

$$\max_{i=1,\ldots,N}\left|\lambda_1 - \lambda_{1,(i)}^{(T_1)}\right| = \mathcal{O}\left(\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}\sqrt{N}|\lambda_N|\lambda_1^2\phi^L\right) = \mathcal{O}\left(\left(\frac{\lambda_2-\lambda_3}{\sqrt{N}}\right)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}\right) = o(\rho).$$

□

Remark 2.2. *The magnitude of eigenvalues in Theorem 2 might depend on N, the number of nodes. For example, if the symmetric matrix $\boldsymbol{Z}$ equals the adjacency matrix $\boldsymbol{A}$ from Stochastic Block Model with 2 communities, then we have $|\lambda_1|, |\lambda_2| = \Theta(\log N)$ and $|\lambda_3| = \mathcal{O}(\sqrt{\log N})$ with high probability (see* **Lemma 3.6***).*

2.2 Centralized estimation error of Theorem 2

In **Lemma 2.2**, we first show

$$\left\|\boldsymbol{G}_c^{(t)}\right\|_2 = \|\lambda_1\boldsymbol{u}_1\boldsymbol{u}_1^{\mathtt{H}}\boldsymbol{u}_2^{(t)} - \gamma_1^{(T_1)}\boldsymbol{q}_1^{(T_1)}(\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)}\|_2 = \mathcal{O}\left(\left(\tan\theta_1^{(0)}\right)|\lambda_1|\left|\frac{\lambda_2}{\lambda_1}\right|^{T_1}\|\boldsymbol{u}_2^{(t)}\|_2\right),$$

and $\|\boldsymbol{u}_2^{(t)}\|_2$ will be bounded in **Lemma 2.10**.

Lemma 2.2. *Define $\boldsymbol{G}_c^{(t)} = \left(\lambda_1\boldsymbol{u}_1\boldsymbol{u}_1^{\mathtt{H}} - \gamma_1^{(t)}\boldsymbol{q}_1^{(T_1)}\left(\boldsymbol{q}_1^{(T_1)}\right)^{\mathtt{H}}\right)\boldsymbol{u}_2^{(t)}$, where T_1 is the number of power iterations for estimating $(\lambda_1, \boldsymbol{u}_1)$. Then*

$$\left\|\boldsymbol{G}_c^{(t)}\right\|_2 = \mathcal{O}\left(\left(\tan\theta_1^{(0)}\right)|\lambda_1|\left|\frac{\lambda_2}{\lambda_1}\right|^{T_1}\|\boldsymbol{u}_2^{(t)}\|_2\right).$$

Proof. Let $\boldsymbol{q}_1^{(T_1)} = \boldsymbol{u}_1 + \boldsymbol{\xi}^{(T_1)}$ and $\gamma_1^{(T_1)} = \lambda_1 + \psi^{(T_1)}$, that is, $\boldsymbol{\xi}^{(T_1)}$ and $\psi^{(T_1)}$ represent the estimation error. We can decompose the perturbation matrix $\boldsymbol{G}_c^{(t)}$ as follows:

$$\begin{aligned}
\boldsymbol{G}_c^{(t)} &= \left(\lambda_1\boldsymbol{u}_1\boldsymbol{u}_1^{\mathtt{H}} - \gamma_1^{(T_1)}\boldsymbol{q}_1^{(T_1)}\left(\boldsymbol{q}_1^{(T_1)}\right)^{\mathtt{H}}\right)\boldsymbol{u}_2^{(t)} \\
&= \left(\lambda_1\boldsymbol{u}_1\boldsymbol{u}_1^{\mathtt{H}} - \left(\lambda_1+\psi^{(T_1)}\right)\left(\boldsymbol{u}_1+\boldsymbol{\xi}^{(T_1)}\right)\left(\boldsymbol{u}_1+\boldsymbol{\xi}^{(T_1)}\right)^{\mathtt{H}}\right)\boldsymbol{u}_2^{(t)} \\
&= \left(\lambda_1\left(\boldsymbol{u}_1(\boldsymbol{\xi}^{(T_1)})^{\mathtt{H}} + \boldsymbol{\xi}^{(T_1)}\boldsymbol{u}_1^{\mathtt{H}} + \boldsymbol{\xi}^{(T_1)}(\boldsymbol{\xi}^{(T_1)})^{\mathtt{H}}\right) + \psi^{(T_1)}\left(\boldsymbol{u}_1\boldsymbol{u}_1^{\mathtt{H}} + \boldsymbol{u}_1(\boldsymbol{\xi}^{(T_1)})^{\mathtt{H}} + \boldsymbol{\xi}^{(T_1)}\boldsymbol{u}_1^{\mathtt{H}} + \boldsymbol{\xi}^{(T_1)}(\boldsymbol{\xi}^{(T_1)})^{\mathtt{H}}\right)\right)\boldsymbol{u}_2^{(t)}.
\end{aligned}$$

Then, we have

$$\begin{aligned}
\left\|\boldsymbol{G}_c^{(t)}\right\|_2 \le\; & |\lambda_1|\left(\left\|\boldsymbol{u}_1(\boldsymbol{\xi}^{(T_1)})^{\mathtt{H}}\right\|_2 + \left\|\boldsymbol{\xi}^{(T_1)}\boldsymbol{u}_1^{\mathtt{H}}\right\|_2 + \left\|\boldsymbol{\xi}^{(T_1)}(\boldsymbol{\xi}^{(T_1)})^{\mathtt{H}}\right\|_2\right)\|\boldsymbol{u}_2^{(t)}\|_2 \\
&+ \left|\psi^{(T_1)}\right|\left(\left\|\boldsymbol{u}_1\boldsymbol{u}_1^{\mathtt{H}}\right\|_2 + \left\|\boldsymbol{u}_1(\boldsymbol{\xi}^{(T_1)})^{\mathtt{H}}\right\|_2 + \left\|\boldsymbol{\xi}^{(T_1)}\boldsymbol{u}_1^{\mathtt{H}}\right\|_2 + \left\|\boldsymbol{\xi}^{(T_1)}(\boldsymbol{\xi}^{(T_1)})^{\mathtt{H}}\right\|_2\right)\|\boldsymbol{u}_2^{(t)}\|_2
\end{aligned}$$

Next, we bound the error terms using **Lemma 4.3**:

$$\|\boldsymbol{\xi}^{(T_1)}\|_2 = \|\boldsymbol{u}_1^{(T_1)} - \boldsymbol{u}_1\|_2 \le \sqrt{2}\tan(\theta_0)\left|\frac{\lambda_2}{\lambda_1}\right|^{T_1}$$

and

$$\left|\psi^{(T_1)}\right| = \left|\gamma_1^{(t)} - \lambda_1\right| =\le \left|\frac{\lambda_1}{\lambda_N}\right|\max_{m=2,\ldots,N}|\lambda_1-\lambda_m|\left(\tan\theta_0\right)^2\left|\frac{\lambda_2}{\lambda_1}\right|^{2t-2}$$

For brevity, let's denote $C=\sqrt{2}\tan\theta_0$ and $D=\left|\frac{\lambda_1}{\lambda_N}\right|\max_{m=2,...,N}|(\lambda_1-\lambda_m)|\left(\tan\theta_0\right)^2$. It is also known that $\|\boldsymbol{\nu}\boldsymbol{v}^{\mathtt{H}}\|_2=\|\boldsymbol{\nu}^{\mathtt{H}}\boldsymbol{v}\|_2$ for any vector $\boldsymbol{\nu},\boldsymbol{v}\in\mathbb{R}^N$. Applying the Cauchy-Schwarz inequality, we have that

$$
\begin{aligned}
\left\|\boldsymbol{u}_1\boldsymbol{u}_1^{\mathtt{H}}\right\|_2 &= 1,\\
\left\|\boldsymbol{u}_1(\boldsymbol{\xi}^{(T_1)})^{\mathtt{H}}\right\|_2 = \left\|(\boldsymbol{\xi}^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_1\right\|_2 &= \left\|\boldsymbol{u}_1^{\mathtt{H}}\boldsymbol{\xi}^{(T_1)}\right\|_2 \le \left\|\boldsymbol{\xi}^{(T_1)}\right\|_2 \le C\left|\frac{\lambda_2}{\lambda_1}\right|^{T_1},\\
\left\|\boldsymbol{\xi}^{(T_1)}(\boldsymbol{\xi}^{(T_1)})^{\mathtt{H}}\right\|_2 = \left\|(\boldsymbol{\xi}^{(T_1)})^{\mathtt{H}}\boldsymbol{\xi}^{(T_1)}\right\|_2 &= \left\|\boldsymbol{\xi}^{(T_1)}\right\|_2^2 \le C^2\left|\frac{\lambda_2}{\lambda_1}\right|^{2T_1}.
\end{aligned}
$$

We can upper bound the perturbation matrix:

$$
\begin{aligned}
\left\|\boldsymbol{G}_c^{(t)}\right\|_2 &\le |\lambda_1|\left(2C\left|\frac{\lambda_2}{\lambda_1}\right|^{T_1}+C^2\left|\frac{\lambda_2}{\lambda_1}\right|^{2T_1}\right)\|\boldsymbol{u}_2^{(t)}\|_2 + D\left|\frac{\lambda_2}{\lambda_1}\right|^{2(T_1-1)}\left(1+2C\left|\frac{\lambda_2}{\lambda_1}\right|^{T_1}+C^2\left|\frac{\lambda_2}{\lambda_1}\right|^{2T_1}\right)\|\boldsymbol{u}_2^{(t)}\|_2\\
&= \mathcal{O}\left(\left(\tan\theta_1^{(0)}\right)|\lambda_1|\left|\frac{\lambda_2}{\lambda_1}\right|^{T_1}\|\boldsymbol{u}_2^{(t)}\|_2\right)
\end{aligned}
$$

□

2.3 Decentralized estimation error of Theorem 2

In this subsection, we analyze $\boldsymbol{G}_d^{(t)}=\gamma_1^{(T_1)}\boldsymbol{q}_1^{(T_1)}(\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)}-\boldsymbol{\Lambda}_{\mathbf{1}}^{(T_1)}\boldsymbol{u}_1^{(T_1)}(\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)}$.

To facilitate analysis, we adopt the use of diagonal matrices to store distributed values at each node. For instance, for any t and $y\in\mathbb{R}$, we define $\boldsymbol{Y}^t=\operatorname{diag}(y_1^t,y_2^t,...,y_N^t)\in\mathbb{R}^{N\times N}$ and $\overline{y}^t=\frac{1}{N}\sum_{i=1}^N y_i^t\in\mathbb{R}$. Similarly, we have $\overline{\boldsymbol{Y}}^t=\overline{y}^t\mathbf{I}_N$, where $\mathbf{I}_N$ denotes the identity matrix of size $N\times N$. Furthermore, the estimated largest eigenvalues distributed among nodes can be represented by the diagonal matrix

$$
\boldsymbol{\Lambda}_1^{(T_1)}=[\lambda_{1,(1)}^{(T_1)},...,\lambda_{1,(N)}^{(T_1)}]\mathbf{I}_N
$$

Let us review the notation. $\boldsymbol{u}_{2,(j)}^{(t)}$ denotes the j_{th} element at the t_{th} itration, $\lambda_{1,(i)}^{(T_1)}$ represents the i_{th} element of the estimated largest eigenvalue obtained after T_1 iterations and $\boldsymbol{u}_{1,(i)}^{(T_1)}$ denotes the i_{th} element of the estimated largest eigenvector obtained after T_1 iterations. Moreover, $b_{12,(i)}^{(t+1)}$ represents the estimation of $(\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)}$ for the node i at iteration $t+1$, where

$$
\left\{b_{sr,(i)}^{(t+1)}\right\}_{i=1}^N=\mathrm{AC}\left(\left\{u_{s,(i)}^{(T_s)}u_{r,(i)}^{(t)}\right\}_{i=1}^N;L\right).
$$

and diagonal matrix $\boldsymbol{B}^{(t+1)}=\operatorname{diag}(b_{12,(1)}^{(t+1)},b_{12,(2)}^{(t+1)},...b_{12,(N)}^{(t+1)})$.

It is important note that the inexact normalization causes the estimation error of the eigenvalue and eigenvector. To analyze the inexactness, we denote i_{th} estimated eigenvector with exact normalization at T_1 iteration as $\boldsymbol{q}_i^{T_1}$ and denote the estimated largest eigenvalue with ideal normalization at T_1 iteration as $\gamma_1^{(T_1)},\boldsymbol{\Gamma}^{(T_1)}=\gamma_1^{(T_1)}\mathbf{I}_N\in\mathbb{R}^{N\times N}$.

By now, $\boldsymbol{G}_d^{(t)}$ can be written as

$$
\begin{aligned}
\boldsymbol{G}_d^{(t)} &= \gamma_1^{(T_1)}\boldsymbol{q}_1^{(T_1)}(\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)}-\boldsymbol{\Lambda}_{\mathbf{1}}^{(T_1)}\boldsymbol{u}_1^{(T_1)}\underbrace{(\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)}}_{\text{Estimated by ``\textbf{Gossip}'' step}}\\
&= \gamma_1^{(T_1)}\boldsymbol{q}_1^{(T_1)}(\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)}-\boldsymbol{\Lambda}_{\mathbf{1}}^{(T_1)}\boldsymbol{u}_1^{(T_1)}N\boldsymbol{B}^{(t+1)}\\
&= \underbrace{\boldsymbol{\Gamma}^{(T_1)}\boldsymbol{q}_1^{(T_1)}(\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)}-\boldsymbol{\Gamma}^{(T_1)}\boldsymbol{q}_1^{(T_1)}N\boldsymbol{B}^{(t+1)}}_{\hat{\boldsymbol{G}}_d^{(t)}}+\underbrace{\left(\boldsymbol{\Gamma}^{(T_1)}\boldsymbol{q}_1^{(T_1)}-\boldsymbol{\Lambda}_{\mathbf{1}}^{(T_1)}\boldsymbol{u}_1^{(T_1)}\right)N\boldsymbol{B}^{(t+1)}}_{\tilde{\boldsymbol{G}}_d^{(t)}}\\
&= \hat{\boldsymbol{G}}_d^{(t)}+\tilde{\boldsymbol{G}}_d^{(t)}.
\end{aligned}
\tag{12}
$$

We denote perturbation matrix related to decentralized computing at t iteration as $\hat{\boldsymbol{G}}_d^{(t)}$ and $\tilde{\boldsymbol{G}}_d^{(t)}$. Moreover, we will show $\|\hat{\boldsymbol{G}}_d^{(t)}\|_2$ and $\|\tilde{\boldsymbol{G}}_d^{(t)}\|_2$ are upper bounded by $\min\{\rho, \frac{1}{\tau\sqrt{N}}\}(|\lambda_2| - |\lambda_3|)$ if the number of gossip iteration L is sufficiently large.

2.3.1 Upper bound for the former term in Equation (12)

We first analyze $\|\hat{\boldsymbol{G}}_d^{(t)}\|_2$, which can be expressed as follows:

$$\begin{aligned}
&\|\boldsymbol{\Gamma}^{(T_1)}\boldsymbol{q}_1^{(T_1)}(\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_{\mathbf{2}}^{(t)} - \boldsymbol{\Gamma}^{(T_1)}\boldsymbol{q}_1^{(T_1)}N\boldsymbol{B}^{(t+1)}\|_2 \\
&\qquad\qquad \leq \|\boldsymbol{\Gamma}^{(T_1)}\|_2\|\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_{\mathbf{2}}^{(t)}\mathbf{I}_N - N\overline{\boldsymbol{B}}^{(t+1)} + N\overline{\boldsymbol{B}}^{(t+1)} - N\boldsymbol{B}^{(t+1)}\|_2 \\
&\qquad\qquad \leq |\lambda_1|\left(\|\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_{\mathbf{2}}^{(t)}\mathbf{I}_N - N\overline{\boldsymbol{B}}^{(t+1)}\|_2 + \|N\overline{\boldsymbol{B}}^{(t+1)} - N\boldsymbol{B}^{(t+1)}\|_2\right),
\end{aligned} \tag{13}$$

where $\|\boldsymbol{\Gamma}^{(T_1)}\|_2 = \gamma_1^{(t)} \leq |\lambda_1|$ by **Lemma 4.3**. Following previous notations, we construct the diagonal matrix as:

$$\overline{\boldsymbol{B}}^{(t+1)} = \operatorname{diag}\left(\overline{b}^{t+1}, \overline{b}^{t+1}, \ldots \overline{b}^{t+1}\right), \text{and } \overline{b}^{t+1} = \frac{1}{N}(\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)}.$$

Therefore, we have

$$\begin{aligned}
&|\lambda_1|\left(\|\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_{\mathbf{2}}^{(t)}\mathbf{I}_N - N\overline{\boldsymbol{B}}^{(t+1)}\|_2 + \|N\overline{\boldsymbol{B}}^{(t+1)} - N\boldsymbol{B}^{(t+1)}\|_2\right) \\
&\quad = |\lambda_1|\left(\|(\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_{\mathbf{2}}^{(t)} - (\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)}\|_2 + \|N\overline{\boldsymbol{B}}^{(t+1)} - N\boldsymbol{B}^{(t+1)}\|_2\right) \\
&\quad \leq |\lambda_1|\left(\|(\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}} - (\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}}\|_2\|\boldsymbol{u}_2^{(t)}\|_2 + \|N\overline{\boldsymbol{B}}^{(t+1)} - N\boldsymbol{B}^{(t+1)}\|_2\right) \\
&\quad \leq |\lambda_1|\left(\|\boldsymbol{q}_1^{(T_1)} - \boldsymbol{u}_1^{(T_1)}\|_2\|\boldsymbol{u}_2^{(t)}\|_2 + \phi^L\|(\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)}\mathbf{I}_N - N\operatorname{diag}\left(\boldsymbol{u}_1^{(T_1)}\right)\operatorname{diag}\left(\boldsymbol{u}_2^{(t)}\right)\|_2\right),
\end{aligned} \tag{14}$$

where the last inequality is due to **Lemma 2.3** and ϕ denotes the second largest eigenvalue of weight matrix $\boldsymbol{W}$.

Lemma 2.3. *([5, Equation (16)]) For any connected graph, suppose agents run **"Gossip"** step in Algorithm 1 with $\boldsymbol{W}$ defined in Equation (1) so that after L gossip iterations, we have*

$$\|\boldsymbol{x}^{(L)} - x_{ave}\mathbf{1}_N\|_2 \leq \phi^L\|\boldsymbol{x}^{(0)} - x_{ave}\mathbf{1}_N\|_2,$$

where ϕ represents the second-largest eigenvalue of $\boldsymbol{W}$.

To obtain the desired upper bound, we first bound the inexactness of local normalization $\|\boldsymbol{q}_1^{(T_1)} - \boldsymbol{u}_1^{(T_1)}\|_2$.

Lemma 2.4. *Let T_1 denote the number of power iteration for estimating $(\lambda_1, \boldsymbol{u}_1)$ and L represent the number of gossip iterations in Algorithm 1. For some constant $C_2 > 0$, we have*

$$\|\tilde{\boldsymbol{u}}_1^{(T_1)} - \boldsymbol{u}_1^{(T_1)}\|_2 \leq C_2\lambda_1^2\|\boldsymbol{u}_1^{(T_1-1)}\|_2^2\,\phi^L,$$

where $\tilde{\boldsymbol{u}}_1^{(T_1)} = \dfrac{\hat{\boldsymbol{u}}_1^{(T_1)}}{\|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2}$ and $\boldsymbol{u}_1^{(T_1)}$ denotes the estimated vector by Algorithm 1.

Proof. Based on the context of

$$\left\{a_{(i)}^{(T_1)}\right\}_{i=1}^N = \mathrm{AC}\left(\left\{\left(\hat{u}_{1,(i)}^{(T_1)}\right)^2\right\}_{i=1}^N; L\right),$$

we define the consensus value as:

$$\forall i, \left(\overline{a}_{(i)}^{(T_1)}\right)^{-1/2} = \left(\overline{a}^{(T_1)}\right)^{-1/2} = \sqrt{\frac{N}{\sum_{j=1}^N\left(\hat{\boldsymbol{u}}_{1,(j)}^{(T_1)}\right)^2}},$$

then $\hat{\boldsymbol{u}}_{1,(i)}^{(T_1)}\left(\overline{a}_{(i)}^{(T_1)}\right)^{-1/2} = \dfrac{\sqrt{N}\hat{\boldsymbol{u}}_{1,(i)}^{(T_1)}}{\sqrt{\sum_{j=1}^{N}\left(\hat{\boldsymbol{u}}_{1,(j)}^{(T_1)}\right)^2}} \leq \sqrt{N}$. Thus, we have

$$
\begin{aligned}
\|\boldsymbol{u}_1^{(T_1)} - \tilde{\boldsymbol{u}}_1^{(T_1)}\|_2 &= \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\left(\hat{\boldsymbol{u}}_{1,(i)}^{(T_1)}\right)\left(\left(a_{(i)}^{(T_1)}\right)^{-1/2} - \left(\overline{a}_{(i)}^{(T_1)}\right)^{-1/2}\right)\Big)^2} \\
&= \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left((\hat{\boldsymbol{u}}_{1,(i)}^{(T_1)})\left(\overline{a}_{(i)}^{(T_1)}\right)^{-1/2}\left(\left(a_{(i)}^{(T_1)}\right)^{1/2} - \left(\overline{a}_{(i)}^{(T_1)}\right)^{1/2}\right)\left(a_{(i)}^{(T_1)}\right)^{-1/2}\right)^2} \\
&\leq \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\sqrt{N}\left(\left(a_{(i)}^{(T_1)}\right)^{1/2} - \left(\overline{a}_{(i)}^{(T_1)}\right)^{1/2}\right)\left(a_{(i)}^{(T_1)}\right)^{-1/2}\right)^2} \\
&= \sqrt{\sum_{i=1}^{N}\left(\left(\left(a_{(i)}^{(T_1)}\right)^{1/2} - \left(\overline{a}_{(i)}^{(T_1)}\right)^{1/2}\right)\left(a_{(i)}^{(T_1)}\right)^{-1/2}\right)^2}
\end{aligned}
\tag{15}
$$

We denote $\left(\min_i\left(a_{(i)}^{(T_1)}\right)^{-1/2}\right)^{-1}$ as $m_1^{(T_1)}$, which is a bounded constant because all $a_{(i)}^{(T_1)}$ are positive constants and would converge to $\overline{a}^{(T_1)}$ as $L \to \infty$. Therefore, we have:

$$
\begin{aligned}
\|\boldsymbol{u}_1^{(T_1)} - \tilde{\boldsymbol{u}}_1^{(T_1)}\|_2 &\leq \frac{1}{\min_i\left(a_{(i)}^{(T_1)}\right)^{-1/2}}\sqrt{\sum_{i=1}^{N}\left(\left(a_{(i)}^{(T_1)}\right)^{1/2} - \left(\overline{a}_{(i)}^{(T_1)}\right)^{1/2}\right)^2} \\
&\leq m_1^{(T_1)}\sqrt{\sum_{i=1}^{N}\left(\left(a_{(i)}^{(T_1)}\right)^{1/2} - \left(\overline{a}_{(i)}^{(T_1)}\right)^{1/2}\right)^2\left(\left(a_{(i)}^{(T_1)}\right)^{1/2} + \left(\overline{a}_{(i)}^{(T_1)}\right)^{1/2}\right)^2} \\
&= m_1^{(T_1)}\left\|\boldsymbol{a}^{(T_1)} - \overline{\boldsymbol{a}}^{(T_1)}\right\|_2 \\
&\leq m_1^{(T_1)}\phi^L\left\|\hat{\boldsymbol{u}}_1^{(T_1)} \circ \hat{\boldsymbol{u}}_1^{(T_1)} - \frac{\|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2^2}{N}\mathbf{1}_N\right\|_2,
\end{aligned}
\tag{16}
$$

where $\circ$ means the element-wise multiplication and the last inequality holds due to the definition of $\overline{\boldsymbol{a}}^{T_1}$ and **Lemma 2.3**. Since $\|\hat{\boldsymbol{u}}_1^{(T_1)} \circ \hat{\boldsymbol{u}}_1^{(T_1)}\|_2 \leq \|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2^2$, we have

$$
\begin{aligned}
\|\hat{\boldsymbol{u}}_1^{(T_1)} \circ \hat{\boldsymbol{u}}_1^{(T_1)} - \frac{\|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2^2}{N}\mathbf{1}_N\|_2 &\leq \|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2^2 + \|\frac{\|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2^2}{N}\mathbf{1}_N\|_2 \\
&= \|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2^2 + \frac{1}{\sqrt{N}}\|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2^2 \\
&\leq 1.5\|\boldsymbol{Z}\boldsymbol{u}_1^{(T_1-1)}\|_2^2 \\
&\leq 1.5\lambda_1^2\|\boldsymbol{u}_1^{(T_1-1)}\|_2^2
\end{aligned}
\tag{17}
$$

Thus for any constant $C_2 \geq 1.5m_1^{(T_1)}$ we have,

$$
\begin{aligned}
\|\tilde{\boldsymbol{u}}_1^{(T_1)} - \boldsymbol{u}_1^{(T_1)}\|_2 &\leq 1.5m_1^{(T_1)}\lambda_1^2\|\boldsymbol{u}_1^{(T_1-1)}\|_2^2\,\phi^L \\
&\leq C_2\lambda_1^2\|\boldsymbol{u}_1^{(T_1-1)}\|_2^2\,\phi^L.
\end{aligned}
\tag{18}
$$

□

Lemma 2.5. *Let T_1 denote the number of power iteration for estimating $(\lambda_1, \boldsymbol{u}_1)$ and L represent the number of gossip iterations in Algorithm 1. We have*

$$
\|\boldsymbol{q}_1^{(T_1)} - \boldsymbol{u}_1^{(T_1)}\|_2 \leq \mathcal{O}\left(\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}\lambda_1^2\phi^L\right), \qquad \max_{i=1,\ldots,N}\left|\lambda_1 - \lambda_{1,(i)}^{(T_1)}\right| = \mathcal{O}\left(\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}\sqrt{N}|\lambda_N|\lambda_1^2\phi^L\right).
$$

where $\tilde{\boldsymbol{u}}_1^{(T_1)} = \frac{\hat{\boldsymbol{u}}_1^{(T_1)}}{\|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2}$, $\overline{\boldsymbol{u}}_1^{(T_1)} = \frac{\hat{\boldsymbol{u}}_1^{(T_1)}}{\|\hat{\boldsymbol{q}}_1^{(T_1)}\|_2}$ *and* $\boldsymbol{u}_1^{(T_1)}$ *denotes the estimated vector by Algorithm 1.*

Proof.

$$\|\boldsymbol{q}_1^{(T_1)} - \boldsymbol{u}_1^{(T_1)}\|_2 \leq \|\boldsymbol{q}_1^{(T_1)} - \overline{\boldsymbol{u}}_1^{(T_1)}\|_2 + \|\overline{\boldsymbol{u}}_1^{(T_1)} - \boldsymbol{u}_1^{(T_1)}\|_2 \tag{19}$$

In Equation (19), the former term can be bounded as follows:

$$\begin{aligned}
\|\boldsymbol{q}_1^{(T_1)} - \overline{\boldsymbol{u}}_1^{(T_1)}\|_2 &= \left\| \frac{\hat{\boldsymbol{q}}_1^{(T_1)}}{\|\hat{\boldsymbol{q}}_1^{(T_1)}\|_2} - \frac{\hat{\boldsymbol{u}}_1^{(T_1)}}{\|\hat{\boldsymbol{q}}_1^{(T_1)}\|_2} \right\|_2 = \frac{1}{\|\hat{\boldsymbol{q}}_1^{(T_1)}\|_2} \left\| \hat{\boldsymbol{q}}_1^{(T_1)} - \hat{\boldsymbol{u}}_1^{(T_1)} \right\|_2 \\
&= \frac{1}{\|\hat{\boldsymbol{q}}_1^{(T_1)}\|_2} \left\| \boldsymbol{Z}\boldsymbol{q}_1^{(T_1-1)} - \boldsymbol{Z}\boldsymbol{u}_1^{(T_1-1)} \right\|_2 \\
&\leq \frac{\|\boldsymbol{Z}\|_2}{\|\hat{\boldsymbol{q}}_1^{(T_1)}\|_2} \left\| \boldsymbol{q}_1^{(T_1-1)} - \boldsymbol{u}_1^{(T_1-1)} \right\|_2 \\
&\leq \left| \frac{\lambda_1}{\lambda_N} \right| \left\| \boldsymbol{q}_1^{(T_1-1)} - \boldsymbol{u}_1^{(T_1-1)} \right\|_2,
\end{aligned} \tag{20}$$

where the last inequality holds due to $|\lambda_1| \geq \|\hat{\boldsymbol{q}}_1^{(T_1)}\|_2 \geq |\lambda_N|$, which is shown in **Lemma 4.3**. In Equation (19), the latter term can be bounded as follows:

$$\begin{aligned}
\|\overline{\boldsymbol{u}}_1^{(T_1)} - \boldsymbol{u}_1^{(T_1)}\|_2 &= \left\| \frac{\hat{\boldsymbol{u}}_1^{(T_1)}}{\|\hat{\boldsymbol{q}}_1^{(T_1)}\|_2} - \frac{\hat{\boldsymbol{u}}_1^{(T_1)}}{\|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2} + \frac{\hat{\boldsymbol{u}}_1^{(T_1)}}{\|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2} - \boldsymbol{u}_1^{(T_1)} \right\|_2 = \frac{1}{\|\hat{\boldsymbol{q}}_1^{(T_1)}\|_2} \left\| \hat{\boldsymbol{q}}_1^{(T_1)} - \hat{\boldsymbol{u}}_1^{(T_1)} \right\|_2 \\
&\leq \left\| \frac{\hat{\boldsymbol{u}}_1^{(T_1)}}{\|\hat{\boldsymbol{q}}_1^{(T_1)}\|_2} - \frac{\hat{\boldsymbol{u}}_1^{(T_1)}}{\|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2} \right\|_2 + \left\| \frac{\hat{\boldsymbol{u}}_1^{(T_1)}}{\|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2} - \boldsymbol{u}_1^{(T_1)} \right\|_2 \\
&\leq \left\| \hat{\boldsymbol{u}}_1^{(T_1)} \right\|_2 \left\| \frac{1}{\|\hat{\boldsymbol{q}}_1^{(T_1)}\|_2} - \frac{1}{\|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2} \right\|_2 + \underbrace{\left\| \frac{\hat{\boldsymbol{u}}_1^{(T_1)}}{\|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2} - \boldsymbol{u}_1^{(T_1)} \right\|_2}_{\textbf{Lemma 2.4} + \textbf{ Lemma 2.9}} \\
&\leq \frac{\left\| \hat{\boldsymbol{u}}_1^{(T_1)} \right\|_2}{\|\hat{\boldsymbol{q}}_1^{(T_1)}\|_2 \|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2} \left\| \|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2 - \|\hat{\boldsymbol{q}}_1^{(T_1)}\|_2 \right\|_2 + C_2 \lambda_1^2 \phi^L \\
&\leq \frac{1}{\|\hat{\boldsymbol{q}}_1^{(T_1)}\|_2} \left\| \hat{\boldsymbol{u}}_1^{(T_1)} - \hat{\boldsymbol{q}}_1^{(T_1)} \right\|_2 + C_2 \lambda_1^2 \phi^L \\
&\leq \frac{\|\boldsymbol{Z}\|_2}{\|\hat{\boldsymbol{q}}_1^{(T_1)}\|_2} \left\| \boldsymbol{q}_1^{(T_1-1)} - \boldsymbol{u}_1^{(T_1-1)} \right\|_2 + C_2 \lambda_1^2 \phi^L \\
&\leq \left| \frac{\lambda_1}{\lambda_N} \right| \left\| \boldsymbol{q}_1^{(T_1-1)} - \boldsymbol{u}_1^{(T_1-1)} \right\|_2 + C_2 \lambda_1^2 \phi^L.
\end{aligned} \tag{21}$$

Finally, we have

$$\begin{aligned}
\|\boldsymbol{q}_1^{(T_1)} - \boldsymbol{u}_1^{(T_1)}\|_2 &\leq \|\boldsymbol{q}_1^{(T_1)} - \overline{\boldsymbol{u}}_1^{(T_1)}\|_2 + \|\overline{\boldsymbol{u}}_1^{(T_1)} - \boldsymbol{u}_1^{(T_1)}\|_2 \\
&\leq 2 \left| \frac{\lambda_1}{\lambda_N} \right| \left\| \boldsymbol{q}_1^{(T_1-1)} - \boldsymbol{u}_1^{(T_1-1)} \right\|_2 + C_2 \lambda_1^2 \phi^L.
\end{aligned} \tag{22}$$

Since $\boldsymbol{q}_1^{(0)}$ and $\boldsymbol{u}_1^{(0)}$ both are initialized by a Gaussian distribution, $\boldsymbol{q}_1^{(0)} = \boldsymbol{u}_1^{(0)}$ and thus we have

$$\|\boldsymbol{q}_1^{(T_1)} - \boldsymbol{u}_1^{(T_1)}\|_2 = \mathcal{O}\left(\left| \frac{\lambda_1}{\lambda_N} \right|^{T_1-1} \lambda_1^2 \phi^L \right).$$

For the estimated eigenvalue, we have

$$\begin{aligned}
\|\lambda_1 \mathbf{I}_N - \boldsymbol{\Lambda}_1^{(T_1)}\|_2 &\leq |\lambda_1 - \gamma_1^{(T_1}| + \|\gamma_1^{(T_1)}\mathbf{I}_N - \boldsymbol{\Lambda}_1^{(T_1)}\|_2 \\
&\leq \underbrace{\left|\frac{\lambda_1}{\lambda_N}\right| \max_{m=2,\ldots,N} |\lambda_1 - \lambda_m| \left(\tan\theta_1^{(0)}\right)^2 \left|\frac{\lambda_2}{\lambda_1}\right|^{2T_1-2}}_{\textbf{Lemma 4.3}} + \max_{i=1,\ldots,N} \left| \|\hat{\boldsymbol{q}}_1^{(T_1)}\|_2 - \|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2 + \|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2 - \lambda_{1,(i)}^{(T_1)} \right| \\
&\leq \left|\frac{\lambda_1}{\lambda_N}\right| \max_{m=2,\ldots,N} |\lambda_1 - \lambda_m| \left(\tan\theta_1^{(0)}\right)^2 \left|\frac{\lambda_2}{\lambda_1}\right|^{2T_1-2} + \left| \|\hat{\boldsymbol{q}}_1^{(T_1)}\|_2 - \|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2 \right| + \max_{i=1,\ldots,N} \left| \|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2 - \lambda_{1,(i)}^{(T_1)} \right| \\
&\leq \left|\frac{\lambda_1}{\lambda_N}\right| \max_{m=2,\ldots,N} |\lambda_1 - \lambda_m| \left(\tan\theta_1^{(0)}\right)^2 \left|\frac{\lambda_2}{\lambda_1}\right|^{2T_1-2} + \lambda_1 \left\| \boldsymbol{q}_1^{(T_1-1)} - \boldsymbol{u}_1^{(T_1-1)} \right\|_2 + \max_{i=1,\ldots,N} \left| \|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2 - \lambda_{1,(i)}^{(T_1)} \right| \\
&= \mathcal{O}\left(\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1} |\lambda_N| \lambda_1^2 \phi^L \right) + \max_{i=1,\ldots,N} \left| \|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2 - \lambda_{1,(i)}^{(T_1)} \right|.
\end{aligned} \tag{23}$$

Based on the context of

$$\left\{ a_{(i)}^{(T_1)} \right\}_{i=1}^N = \mathrm{AC}\left(\left\{ \left(\hat{u}_{1,(i)}^{(T_1)} \right)^2 \right\}_{i=1}^N ; L \right),$$

we define the consensus value as:

$$\forall i, \overline{a}_{(i)}^{(T_1)} = \overline{a}^{(T_1)} = \sqrt{\frac{\sum_{j=1}^N \left(\hat{\boldsymbol{u}}_{1,(j)}^{(T_1)} \right)^2}{N}},$$

and $\lambda_1^{(T_1)} = \sqrt{N a_{(i)}^{(T_1)}}$. Similar to Equation (17), we have

$$\begin{aligned}
\max_{i=1,\ldots,N} \left| \|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2 - \lambda_{1,(i)}^{(T_1)} \right| &= \left| \sqrt{N \overline{a}^{(T_1)}} - \sqrt{N a_{(i)}^{(T_1)}} \right| \\
&\leq \sqrt{N} \|\boldsymbol{a}^{(T_1)} - \overline{\boldsymbol{a}}^{(T_1)}\|_2 \\
&\leq \sqrt{N} \phi^L \left\| \hat{\boldsymbol{u}}_1^{(T_1)} \circ \hat{\boldsymbol{u}}_1^{(T_1)} - \frac{\|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2^2}{N} \mathbf{1}_N \right\|_2 \\
&= \mathcal{O}\left(\sqrt{N} \lambda_1^2 \phi^L \right).
\end{aligned} \tag{24}$$

Finally, we obtain

$$\begin{aligned}
\|\lambda_1 \mathbf{I}_N - \boldsymbol{\Lambda}_1^{(T_1)}\|_2 &= \max_{i=1,\ldots,N} \left| \lambda_1 - \lambda_{1,(i)}^{(T_1)} \right| \\
&= \mathcal{O}\left(\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1} |\lambda_N| \lambda_1^2 \phi^L \right) + \max_{i=1,\ldots,N} \left| \|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2 - \lambda_{1,(i)}^{(T_1)} \right| \\
&= \mathcal{O}\left(\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1} |\lambda_N| \lambda_1^2 \phi^L \right) + \mathcal{O}\left(\sqrt{N} \lambda_1^2 \phi^L \right) \\
&= \mathcal{O}\left(\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1} \sqrt{N} |\lambda_N| \lambda_1^2 \phi^L \right)
\end{aligned} \tag{25}$$

□

Next, we show $\left\| (\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}} \boldsymbol{u}_2^{(t)} \mathbf{I}_N - N \mathrm{diag}\left(\boldsymbol{u}_1^{(T_1)} \right) \mathrm{diag}\left(\boldsymbol{u}_2^{(t)} \right) \right\|_2$ in Equation (14) is bounded.

Lemma 2.6. *At each node i, the maximal distance between initial value $u_{1,(i)}^{(T_1)} u_{2,(i)}^{(t)}$ and the average $\frac{1}{N} \boldsymbol{u}_1^{(T_1)})^{\mathtt{H}} \boldsymbol{u}_2^{(t)}$ can be expressed as:*

$$\left\| (\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}} \boldsymbol{u}_2^{(t)} \mathbf{I}_N - N \, diag\left(\boldsymbol{u}_1^{(T_1)} \right) diag\left(\boldsymbol{u}_2^{(t)} \right) \right\|_2 \leq \left\| \boldsymbol{u}_1^{(T_1)} \right\|_2 \left\| \boldsymbol{u}_2^{(t)} \right\|_2 (1+N).$$

Note that the operator 2-norm of diagonal matrix equals the absolute value of entry with largest magnitude.

Proof.

$$
\begin{aligned}
\left\|\operatorname{diag}\left((\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)}\right) - N\operatorname{diag}\left(\boldsymbol{u}_1^{(T_1)}\right)\operatorname{diag}\left(\boldsymbol{u}_2^{(t)}\right)\right\|_2 &\le \left\|\boldsymbol{u}_1^{(T_1)}\right\|_2\left\|\boldsymbol{u}_2^{(t)}\right\|_2 + \left\|N\operatorname{diag}\left(\boldsymbol{u}_1^{(T_1)}\right)\operatorname{diag}\left(\boldsymbol{u}_2^{(t)}\right)\right\|_2 \\
&= \left\|\boldsymbol{u}_1^{(T_1)}\right\|_2\left\|\boldsymbol{u}_2^{(t)}\right\|_2 + N\max_{i=1,\dots,N}\left|(\boldsymbol{u}_1^{(T_1)})_i(\boldsymbol{u}_2^{(t)})_i\right| \\
&\le \left\|\boldsymbol{u}_1^{(T_1)}\right\|_2\left\|\boldsymbol{u}_2^{(t)}\right\|_2 + N\left\|\boldsymbol{u}_1^{(T_1)}\right\|_2\left\|\boldsymbol{u}_2^{(t)}\right\|_2 \\
&= \left\|\boldsymbol{u}_1^{(T_1)}\right\|_2\left\|\boldsymbol{u}_2^{(t)}\right\|_2(1+N).
\end{aligned} \tag{26}
$$

□

Finally, we can derive the upper bound for $\|\hat{\boldsymbol{G}}_d^{(t)}\|_2$, as described in **Lemma 2.7**.

Lemma 2.7. *Given* $\|\hat{\boldsymbol{G}}_d^{(t)}\|_2 = \left\|\boldsymbol{\Gamma}^{t+1}\boldsymbol{q}_1^{(T_1)}(\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)} - \boldsymbol{\Gamma}^{t+1}\boldsymbol{q}_1^{(T_1)}N\boldsymbol{B}^{(t+1)}\right\|_2$, *we have*

$$
\left\|\hat{\boldsymbol{G}}_d^{(t)}\right\|_2 \le \lambda_1\left\|\boldsymbol{u}_2^{(t)}\right\|_2\phi^L\left(C_3\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}\lambda_1^2\phi^L + (1+N)\left\|\boldsymbol{u}_1^{(T_1)}\right\|_2\right),
$$

for some constant $C_3 > 0$.

Proof. Combine Lemma 2.5 and Lemma 2.6, we can derive

$$
\begin{aligned}
\|\hat{\boldsymbol{G}}_d^{(t)}\|_2 &= \left\|\boldsymbol{\Gamma}^{t+1}\boldsymbol{q}_1^{(T_1)}(\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)} - \boldsymbol{\Gamma}^{t+1}\boldsymbol{q}_1^{(T_1)}N\boldsymbol{B}^{(t+1)}\right\|_2 \\
&\le \lambda_1\left(\left\|\boldsymbol{q}_1^{(T_1)} - \boldsymbol{u}_1^{(T_1)}\right\|_2\left\|\boldsymbol{u}_2^{(t)}\right\|_2 + \phi^L\left\|\operatorname{diag}\left((\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)}\right) - N\operatorname{diag}\left(\boldsymbol{u}_1^{(T_1)}\right)\operatorname{diag}\left(\boldsymbol{u}_2^{(t)}\right)\right\|_2\right) \\
&\le \lambda_1\left(C_3\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}\lambda_1^2\phi^L\left\|\boldsymbol{u}_2^{(t)}\right\|_2 + (1+N)\left\|\boldsymbol{u}_1^{(T_1-1)}\right\|_2\left\|\boldsymbol{u}_2^{(t)}\right\|_2\phi^L\right) \\
&\le \lambda_1\left\|\boldsymbol{u}_2^{(t)}\right\|_2\phi^L\left(C_3\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}\lambda_1^2 + (1+N)\left\|\boldsymbol{u}_1^{(T_1)}\right\|_2\right).
\end{aligned} \tag{27}
$$

□

2.3.2 Upper bound for the latter term in Equation (12)

Now, we analyze $\left\|\tilde{\boldsymbol{G}}_d^{(t)}\right\|_2 = \left\|\left(\boldsymbol{\Gamma}^{(T_1)}\boldsymbol{q}_1^{(T_1)} - \boldsymbol{\Lambda}_1^{(T_1)}\boldsymbol{u}_1^{(T_1)}\right)N\boldsymbol{B}^{(t+1)}\right\|_2$ in **Lemma 2.8**.

Lemma 2.8. *Given* $\left\|\tilde{\boldsymbol{G}}_d^{(t)}\right\|_2 = \left\|\left(\boldsymbol{\Gamma}^{(T_1)}\boldsymbol{q}_1^{(T_1)} - \boldsymbol{\Lambda}_1^{(T_1)}\boldsymbol{u}_1^{(T_1)}\right)N\boldsymbol{B}^{(t+1)}\right\|_2$, *we have*

$$
\left\|\tilde{\boldsymbol{G}}_d^{(t)}\right\|_2 \le C_3N|\lambda_1^3|\left\|\boldsymbol{u}_1^{(T_1)}\right\|_2\left\|\boldsymbol{u}_2^{(t)}\right\|_2\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}\phi^L.
$$

for some constant $C_2 > 0$.

Proof. We have

$$
\begin{aligned}
\left\|\left(\boldsymbol{\Gamma}^{(T_1)}\boldsymbol{q}_1^{(T_1)} - \boldsymbol{\Lambda}_1^{(T_1)}\boldsymbol{u}_1^{(T_1)}\right)N\boldsymbol{B}^{(t+1)}\right\|_2 &\le \left\|N\boldsymbol{B}^{(t+1)}\right\|_2\left\|\boldsymbol{\Gamma}^{(T_1)}\boldsymbol{q}_1^{(T_1)} - \boldsymbol{\Lambda}_1^{(T_1)}\boldsymbol{u}_1^{(T_1)}\right\|_2 \\
&\le N\left\|\boldsymbol{B}^{(t+1)}\right\|_2\left\|\boldsymbol{Z}\boldsymbol{q}_1^{(T_1-1)} - \boldsymbol{Z}\boldsymbol{u}_1^{(T_1)}\right\|_2 \\
&\le N\left\|\boldsymbol{B}^{(t+1)}\right\|_2\|\boldsymbol{Z}\|_2\left\|\boldsymbol{q}_1^{(T_1-1)} - \boldsymbol{u}_1^{(T_1-1)}\right\|_2,
\end{aligned} \tag{28}
$$

where $\boldsymbol{B}^{(t+1)}$ is a diagonal matrix and $\|\boldsymbol{Z}\|_2 = \lambda_1$. Recall **Lemma 2.5**, for some $C_3 > 0$,

$$\left\|\boldsymbol{q}_1^{(T_1-1)} - \boldsymbol{u}_1^{(T_1-1)}\right\|_2 \leq C_3 \left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1} \lambda_1^2\, \phi^L. \tag{29}$$

On the other hand, the output of **"Gossip"** step forms the diagonal matrix $\boldsymbol{B}^{(t+1)} = \operatorname{diag}\left(b_{12,(1)}^{(t+1)}, b_{12,(2)}^{(t+1)}, \ldots b_{12,(N)}^{(t+1)}\right)$. In **Algorithm 1**, $\forall i, b_{12,(i)}^{(t+1)}$ achieve the average as the number of gossip iteration increases, that is,

$$\forall i, b_{12,(i)}^{(t+1)} \leq \max_{i=1,\ldots,N} \left|(\boldsymbol{u}_1^{(T_1)})_i(\boldsymbol{u}_2^{(t)})_i\right|,$$

where $\max_{i=1,\ldots,N} \left|(\boldsymbol{u}_1^{(T_1)})_i(\boldsymbol{u}_2^{(t)})_i\right|$ denotes the absolute value of initial value with largest magnitude. Therefore, we can derive

$$\begin{aligned}
N\|\boldsymbol{Z}\|_2 \left\|\boldsymbol{B}^{(t+1)}\right\|_2 \left\|\boldsymbol{q}_1^{(T_1-1)} - \boldsymbol{u}_1^{(T_1-1)}\right\|_2 &\leq N\lambda_1 \max_{i=1,\ldots,N} \left|(\boldsymbol{u}_1^{(T_1)})_i(\boldsymbol{u}_2^{(t)})_i\right| \left\|\boldsymbol{q}_1^{(T_1-1)} - \boldsymbol{u}_1^{(T_1-1)}\right\|_2 \\
&\leq N\lambda_1 \left\|\boldsymbol{u}_1^{(T_1)}\right\|_2 \left\|\boldsymbol{u}_2^{(t)}\right\|_2 \left\|\boldsymbol{q}_1^{(T_1-1)} - \boldsymbol{u}_1^{(T_1-1)}\right\|_2 \\
&\overset{(29)}{\leq} N\lambda_1 \left\|\boldsymbol{u}_1^{(T_1)}\right\|_2 \left\|\boldsymbol{u}_2^{(t)}\right\|_2 C_3 \left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1} \lambda_1^2\, \phi^L \\
&\leq C_3 N|\lambda_1^3| \left\|\boldsymbol{u}_1^{(T_1)}\right\|_2 \left\|\boldsymbol{u}_2^{(t)}\right\|_2 \left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1} \phi^L.
\end{aligned} \tag{30}$$

In other words, we have

$$\left\|\hat{\boldsymbol{G}}_d^{(t)}\right\|_2 \leq C_3 N|\lambda_1^3| \left\|\boldsymbol{u}_1^{(T_1)}\right\|_2 \left\|\boldsymbol{u}_2^{(t)}\right\|_2 \left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1} \phi^L. \tag{31}$$

□

2.3.3 Upper bound for the length of estimated eigenvectors

Observing **Lemma 2.7** and **Lemma 2.8**, what remains to bound are

$$\forall t, \|\boldsymbol{u}_1^{(t)}\|_2, \|\boldsymbol{u}_2^{(t)}\|_2.$$

At the initialization, we independently generate $\boldsymbol{u}_{1,(i)}^{(0)}, \boldsymbol{u}_{2,(i)}^{(0)} \sim \mathcal{N}(0, \frac{1}{N}), \forall i \in [N]$ such that $\|\boldsymbol{u}_1^{(0)}\|_2, \|\boldsymbol{u}_2^{(0)}\|_2 < \frac{3}{2}$ with probability at least $1 - e^{-\left(\frac{5}{8} - \log\frac{3}{2}\right)N}$, as presented in **Lemma 4.1**. Therefore, we will show $\forall t, |\boldsymbol{u}_1^{(t)}|_2, |\boldsymbol{u}_2^{(t)}|_2 \leq \frac{3}{2}$ by mathematical induction.

Lemma 2.9. *Let T_1 denote the number of power iteration for estimating $(\lambda_1, \boldsymbol{u}_1)$ and L represent the number of gossip iterations in Algorithm 1. Given $\|\boldsymbol{u}_1^{(0)}\|_2 \leq \frac{3}{2}$. For $\tau > 1$, suppose $L = \Omega\left(\frac{1}{\log(\phi^{-1})} \log\left(\frac{\lambda_1^3 N}{(\lambda_2 - \lambda_3)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}}\right)\right)$, we have*

$$\|\boldsymbol{u}_1^{(t)}\|_2 \leq \frac{3}{2}, \forall t.$$

where $\boldsymbol{u}_1^{(T_1)}$ denotes the estimated vector by Algorithm 1.

Proof. For $t \geq 1$, we first assume $\|\boldsymbol{u}_1^{(t-1)}\|_2 \leq \frac{3}{2}$. Subsequently, using **Lemma 2.4**, we have

$$\begin{aligned}
\|\boldsymbol{u}_1^{(t)}\|_2 = \|\tilde{\boldsymbol{u}}_1^{(t)} + \boldsymbol{u}_1^{(t)} - \tilde{\boldsymbol{u}}_1^{(t)}\|_2 &\leq 1 + \|\tilde{\boldsymbol{u}}_1^{(t)} - \boldsymbol{u}_1^{(t)}\|_2 \\
&\leq 1 + C_2\lambda_1^2 \|\boldsymbol{u}_1^{(t-1)}\|_2^2\, \phi^L.
\end{aligned} \tag{32}$$

Here, $\tilde{\boldsymbol{u}}_1^{(t)} = \frac{\hat{\boldsymbol{u}}_1^{(T_1)}}{\|\hat{\boldsymbol{u}}_1^{(T_1)}\|_2}$. For $\tau > 1$, suppose $L = \Omega\left(\frac{1}{\log(\phi^{-1})}\log\left(\frac{\lambda_1^3 N}{(\lambda_2-\lambda_3)\min\{\rho,\frac{1}{\tau\sqrt{N}}\}}\right)\right)$. In conjunction with the mathematical induction, for some sufficiently large N, we have

$$\begin{aligned}
C_2\lambda_1^2\|\boldsymbol{u}_1^{(t-1)}\|_2^2\phi^L &\leq C_2\lambda_2^2\frac{9}{4}\lambda_1^{-3}N^{-1}(\lambda_2-\lambda_3)\ \min\{\rho,\frac{1}{\tau\sqrt{N}}\} \\
&\leq C_2\frac{9}{4}\lambda_1^2\lambda_1^{-3}N^{-1}\ \lambda_1\ \min\{\rho,\frac{1}{\tau\sqrt{N}}\} \\
&\leq \mathcal{O}\left(\min\{\rho,\frac{1}{\tau\sqrt{N}}\}N^{-1}\right) \\
&\leq \frac{1}{2}
\end{aligned} \tag{33}$$

Therefore, we have $\forall t$, $\left\|\boldsymbol{u}_1^{(t)}\right\|_2 \leq \frac{3}{2}$ by the mathematical induction. $\square$

Now, we show $\forall t$, $\left\|\boldsymbol{u}_2^{(t)}\right\|_2 \leq \frac{3}{2}$. Recall $\|\boldsymbol{u}_2^{(0)}\|_2 \leq \frac{3}{2}$ with high probability and we have

$$\|\boldsymbol{u}_2^{(t)}\|_2 = \|\tilde{\boldsymbol{u}}_2^{(t)} + \boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)}\|_2 \leq 1 + \|\boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)}\|_2. \tag{34}$$

Similar to **Lemma 2.4**, we can bound $\|\boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)}\|_2$ as follows:

Lemma 2.10. *Let T_2 denote the number of power iteration for estimating $(\lambda_2, \boldsymbol{u}_2)$ and L represent the number of gossip iterations in Algorithm 1. Given $\|\boldsymbol{u}_2^{(0)}\|_2 \leq \frac{3}{2}$. For $\tau > 1$, suppose $L = \Omega\left(\frac{1}{\log(\phi^{-1})}\log\left(\frac{\lambda_1^3 N}{(\lambda_2-\lambda_3)\min\{\rho,\frac{1}{\tau\sqrt{N}}\}}\right)\right)$, we have*

$$\|\tilde{\boldsymbol{u}}_2^{(T_2)} - \boldsymbol{u}_2^{(T_2)}\|_2 \leq C_2\lambda_1^2\phi^L = \mathcal{O}\left(\frac{(\lambda_2-\lambda_3)\min\{\rho,\frac{1}{\tau\sqrt{N}}\}}{N\lambda_1}\right),\ \textit{and}\ \ \|\boldsymbol{u}_2^{(t)}\|_2 \leq \frac{3}{2}, \forall t.$$

where $\tilde{\boldsymbol{u}}_2^{(T_2)} = \frac{\hat{\boldsymbol{u}}_2^{(T_2-1)}}{\|\hat{\boldsymbol{u}}_2^{(T_2-1)}\|_2}$, $\hat{\boldsymbol{u}}_2^{(t+1)} = \left(\boldsymbol{Z} - \boldsymbol{\Lambda}_1^{(T_1)}\boldsymbol{u}_1^{(T_1)}(\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}}\right)\boldsymbol{u}_2^{(t)}$ and $\boldsymbol{u}_2^{(T_2)}$ denotes the estimated vector by Algorithm 1.

Proof. Based on the context of

$$\left\{c_{(i)}^{(t+1)}\right\}_{i=1}^N = \mathrm{AC}\left(\left\{\left(\hat{u}_{2,(i)}^{(t+1)}\right)^2\right\}_{i=1}^N ; L\right),$$

we define the consensus value as:

$$\forall i, \left(\bar{c}_i^{(t+1)}\right)^{-1/2} = \left(\bar{c}^{(t+1)}\right)^{-1/2} = \sqrt{\frac{N}{\sum_{j=1}^N\left(\hat{\boldsymbol{u}}_{2,(j)}^{(t+1)}\right)^2}}, \left(\bar{\boldsymbol{c}}^{(t+1)}\right)^{-1/2} = \left(\bar{c}^{(t+1)}\right)^{-1/2}\mathbf{1}_N.$$

Similar to Equation (17), we have

$$\begin{aligned}
\|\boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)}\|_2 &\leq m_2^{(t)}\phi^L\|\hat{\boldsymbol{u}}_2^{(t)} \circ \hat{\boldsymbol{u}}_2^{(t)} - \frac{\|\hat{\boldsymbol{u}}_2^{(t)}\|_2^2}{N}\mathbf{1}_N\|_2 \\
&\leq m_2^{(t)}\phi^L\frac{3}{2}\|\hat{\boldsymbol{u}}_2^{(t)}\|_2^2 \\
&= m_2^{(t)}\phi^L\frac{3}{2}\|\boldsymbol{Z}\boldsymbol{u}_2^{(t-1)} - \boldsymbol{\Lambda}_1^{(T_1)}\boldsymbol{u}_1^{T_1}N\boldsymbol{B}^{(t)}\|_2^2 \\
&= m_2^{(t)}\phi^L\frac{3}{2}\|\boldsymbol{Z}\boldsymbol{u}_2^{(t-1)} - \boldsymbol{\Gamma}^{(T_1)}\boldsymbol{q}_1^{(T_1)}(\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t-1)} + \boldsymbol{\Gamma}^{(T_1)}\boldsymbol{q}_1^{(T_1)}(\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t-1)} - \boldsymbol{\Lambda}_1^{(T_1)}\boldsymbol{u}_1^{T_1}N\boldsymbol{B}^{(t)}\|_2^2 \\
&= m_2^{(t)}\phi^L\frac{3}{2}\|\boldsymbol{Z}\boldsymbol{u}_2^{(t-1)} - \boldsymbol{\Gamma}^{(T_1)}\boldsymbol{q}_1^{(T_1)}(\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t-1)} + \hat{\boldsymbol{G}}_d^{(t-1)} + \tilde{\boldsymbol{G}}_d^{(t-1)}\|_2^2
\end{aligned} \tag{35}$$

where the constant $m_2^{(t)} = (\min_i \left(c_i^{(t)}\right)^{-1/2})^{-1}$ and $\hat{\boldsymbol{G}}_d^{(t-1)}, \tilde{\boldsymbol{G}}_d^{(t-1)}$ are defined in Equation (12). Then, we have

$$
\begin{aligned}
\|\boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)}\|_2 &\leq m_2^{(t)} \phi^L \frac{3}{2} \|\boldsymbol{Z}\boldsymbol{u}_2^{(t-1)} - \boldsymbol{\Gamma}^{(T_1)} \boldsymbol{q}_1^{(T_1)} (\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}} \boldsymbol{u}_2^{(t-1)} + \hat{\boldsymbol{G}}_d^{(t-1)} + \tilde{\boldsymbol{G}}_d^{(t-1)}\|_2^2 \\
&\leq m_2^{(t)} \phi^L \frac{3}{2} \left(|\lambda_1| \|\boldsymbol{u}_2^{(t-1)}\|_2 + \gamma_1^{(T_1)} \|\boldsymbol{u}_2^{(t-1)}\|_2 + \|\hat{\boldsymbol{G}}_d^{(t-1)}\|_2 + \|\tilde{\boldsymbol{G}}_d^{(t-1)}\|_2 \right)^2 \\
&\leq m_2^{(t)} \phi^L \frac{3}{2} \left(\frac{3}{2} |\lambda_1| + \frac{3}{2} |\lambda_1| + \|\hat{\boldsymbol{G}}_d^{(t-1)}\|_2 + \|\tilde{\boldsymbol{G}}_d^{(t-1)}\|_2 \right)^2 \\
&\leq m_2^{(t)} \phi^L \frac{3}{2} \left(3|\lambda_1| + \|\hat{\boldsymbol{G}}_d^{(t-1)}\|_2 + \|\tilde{\boldsymbol{G}}_d^{(t-1)}\|_2 \right)^2
\end{aligned}
\tag{36}
$$

By **lemma 2.7** and **lemma 2.8**, we have

$$
\begin{aligned}
\left\|\hat{\boldsymbol{G}}_d^{(t-1)}\right\|_2 &\leq \lambda_1 \left\|\boldsymbol{u}_2^{(t-1)}\right\|_2 \phi^L \left(C_3 \left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1} \lambda_1^2 \phi^L + (1+N) \left\|\boldsymbol{u}_1^{(T_1)}\right\|_2 \right) \\
&\leq \left\|\boldsymbol{u}_2^{(t-1)}\right\|_2 \lambda_1 \phi^L \left(C_3 \left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1} \lambda_1^2 \phi^L + \frac{3}{2}(1+N) \right), \\
\left\|\tilde{\boldsymbol{G}}_d^{(t-1)}\right\|_2 &\leq C_3 N |\lambda_1^3| \left\|\boldsymbol{u}_1^{(T_1)}\right\|_2 \left\|\boldsymbol{u}_2^{(t-1)}\right\|_2 \left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1} \phi^L \\
&\leq C_3 N |\lambda_1^3| \frac{3}{2} \phi^L \left\|\boldsymbol{u}_2^{(t-1)}\right\|_2 \left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}
\end{aligned}
$$

Substituting into Equation (36), we have

$$
\begin{aligned}
\|\boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)}\|_2 &\leq m_2^{(t)} \phi^L \frac{3}{2} \left(3|\lambda_1| + \left\|\boldsymbol{u}_2^{(t-1)}\right\|_2 \lambda_1 \phi^L \left(C_3 \left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1} \lambda_1^2 \phi^L + \frac{3}{2}(1+N) \right) + C_3 N |\lambda_1^3| \frac{3}{2} \phi^L \left\|\boldsymbol{u}_2^{(t-1)}\right\|_2 \left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1} \right)^2 \\
&= m_2^{(t)} \phi^L \frac{3}{2} \left(3|\lambda_1| + \left\|\boldsymbol{u}_2^{(t-1)}\right\|_2 |\lambda_1^3| \phi^{2L} C_3 \left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1} + \left\|\boldsymbol{u}_2^{(t-1)}\right\|_2 \lambda_1 \phi^L \frac{3}{2}(1+N) + C_4 N |\lambda_1^3| \phi^L \left\|\boldsymbol{u}_2^{(t-1)}\right\|_2 \left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1} \right)^2
\end{aligned}
$$

For $\tau > 1$, suppose

$$
L = \Omega\left(\frac{1}{\log(\phi^{-1})} \log\left(\frac{\lambda_1^4 N \left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}}{(\lambda_2 - \lambda_3)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}} \right) \right).
$$

By the mathematical induction, we first assume $\|\boldsymbol{u}_2^{(t-1)}\|_2 \leq \frac{3}{2}$, and thus we have

$$
\|\boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)}\|_2 = \mathcal{O}\left(\phi^{3L} \lambda_1^4 \left\|\boldsymbol{u}_2^{(t-1)}\right\|_2^2 \left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1} \right) = \mathcal{O}\left(\frac{(\lambda_2 - \lambda_3)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}}{N} \right) \leq \frac{1}{2}. \tag{37}
$$

That is to say, for some sufficiently large N, we have

$$
\begin{aligned}
\left\|\boldsymbol{u}_2^{(t)}\right\|_2 &\leq \|\tilde{\boldsymbol{u}}_2^{(t)}\|_2 + \|\tilde{\boldsymbol{u}}_2^{(t)} - \boldsymbol{u}_2^{(t)}\|_2 \\
&\leq 1 + \mathcal{O}\left(\frac{(\lambda_2 - \lambda_3)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}}{N} \right) \leq \frac{3}{2}.
\end{aligned}
$$

The proof is completed the by induction.

□

2.4 Proof of Theorem 2

Now, we can analyze the term $\|\boldsymbol{G}_{cd}^{(t)}\|_2 = \|\boldsymbol{Z} - \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^{\mathtt{H}} \left(\boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)}\right)\|_2, \|\boldsymbol{G}_c^{(t)}\|_2$ and $\|\boldsymbol{G}_d^{(t)}\|_2$.

Lemma 2.11. *Given* $\boldsymbol{G}_d^{(t)} = \boldsymbol{\Lambda}_1^{(t)}\boldsymbol{u}_1^{(t)}N\boldsymbol{B}^{(t)} - \gamma_1^{(T_1)}\boldsymbol{q}_1^{(T_1)}(\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t-1)}$, *using* ***lemma* 2.7** *and* ***lemma* 2.8**, *we have*

$$\|\boldsymbol{G}_d^{(t)}\|_2 \leq \left\|\hat{\boldsymbol{G}}_d^{(t-1)}\right\|_2 + \left\|\tilde{\boldsymbol{G}}_d^{(t-1)}\right\|_2 = \mathcal{O}\left(N|\lambda_1^3|\phi^L\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}\right).$$

Proof.

$$\begin{aligned}
\left\|\hat{\boldsymbol{G}}_d^{(t-1)}\right\|_2 &\leq \lambda_1\left\|\boldsymbol{u}_2^{(t-1)}\right\|_2\phi^L\left(C_3\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}\lambda_1^2\phi^L + (1+N)\left\|\boldsymbol{u}_1^{(T_1)}\right\|_2\right)\\
&\leq \frac{3}{2}\lambda_1\phi^L\left(C_3\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}\lambda_1^2\phi^L + \frac{3}{2}(1+N)\right),\\
\left\|\tilde{\boldsymbol{G}}_d^{(t-1)}\right\|_2 &\leq C_3N|\lambda_1^3|\left\|\boldsymbol{u}_1^{(T_1)}\right\|_2\left\|\boldsymbol{u}_2^{(t-1)}\right\|_2\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}\phi^L\\
&\leq C_3\frac{3}{2}N|\lambda_1^3|\frac{3}{2}\phi^L\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}
\end{aligned}$$

□

Ultimately, we can obtain the desired result in the following Lemma.

Lemma 2.12. *Given*

$$\begin{aligned}
\boldsymbol{G}_{cd}^{(t)} &= \boldsymbol{Z} - \lambda_1\boldsymbol{u}_1\boldsymbol{u}_1^{\mathtt{H}}\left(\boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)}\right)\\
\boldsymbol{G}_c^{(t)} &= \lambda_1\boldsymbol{u}_1\boldsymbol{u}_1^{\mathtt{H}}\boldsymbol{u}_2^{(t)} - \gamma_1^{(T_1)}\boldsymbol{q}_1^{(T_1)}(\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)}\\
\boldsymbol{G}_d^{(t)} &= \gamma_1^{(T_1)}\boldsymbol{q}_1^{(T_1)}(\boldsymbol{q}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)} - \boldsymbol{\Lambda}_1^{(T_1)}\boldsymbol{u}_1^{(T_1)}(\boldsymbol{u}_1^{(T_1)})^{\mathtt{H}}\boldsymbol{u}_2^{(t)}.
\end{aligned}$$

We have

$$\left\|\boldsymbol{G}^{(t)}\right\|_2 \leq \left\|\boldsymbol{G}_{cd}^{(t)}\right\|_2 + \left\|\boldsymbol{G}_c^{(t)}\right\|_2 + \left\|\boldsymbol{G}_d^{(t)}\right\|_2 \leq \mathcal{O}\left((\lambda_2-\lambda_3)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}\right)$$

when the number of power iterations

$$T_1 = \Omega\left(\frac{\lambda_2}{\log\left(\frac{\lambda_1}{\lambda_2}\right)(\lambda_2-\lambda_3)}\log\left(\frac{\tau\sqrt{N}\lambda_2\left|\frac{\lambda_1}{\lambda_N}\right|}{(\lambda_2-\lambda_3)\min\{\rho,\frac{1}{\tau\sqrt{N}}\}}\right)\right),\ T_2 = \Omega\left(\frac{\lambda_2}{\lambda_2-\lambda_3}\log\left(\frac{\tau\sqrt{N}}{\rho}\right)\right).$$

and the number of gossip iterations

$$L = \Omega\left(\frac{1}{\log\left(\phi^{-1}\right)}\frac{\lambda_2}{\log\left(\frac{\lambda_1}{\lambda_2}\right)(\lambda_2-\lambda_3)}\log\left(\frac{\tau N\lambda_2\left|\frac{\lambda_1}{\lambda_N}\right|^2}{(\lambda_2-\lambda_3)\min\{\rho,\frac{1}{\tau\sqrt{N}}\}}\right)\right),$$

for $\tau > 1$.

Proof. Combining **lemma 2.10**, **lemma 2.2** and **lemma 2.11**, we have

$$\begin{aligned}
\left\|\boldsymbol{G}_{cd}^{(T_2)}\right\|_2 &\leq |\lambda_2|\left\|\boldsymbol{u}_2^{(t)} - \tilde{\boldsymbol{u}}_2^{(t)}\right\|_2 \leq C_2\lambda_1^2|\lambda_2|\phi^L\\
\|\boldsymbol{G}_c^{(t)}\|_2 &= \mathcal{O}\left(\left(\tan\theta_1^{(0)}\right)|\lambda_1|\left|\frac{\lambda_2}{\lambda_1}\right|^{T_1}\right),\\
\|\boldsymbol{G}_d^{(t)}\|_2 &= \mathcal{O}\left(N|\lambda_1^3|\phi^L\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1-1}\right).
\end{aligned}$$

To apply **Theorem 1**, $\left(\tan\theta_1^{(0)}\right)|\lambda_1|\left|\frac{\lambda_2}{\lambda_1}\right|^{T_1}$ should less than $(\lambda_2-\lambda_3)\min\{\rho,\frac{1}{\tau\sqrt{N}}\}$, which means

$$T_1 \geq \frac{1}{\log(\lambda_1)-\log(\lambda_2)}\log\left(\frac{\tau\sqrt{N}|\lambda_1|}{(\lambda_2-\lambda_3)\min\{\rho,\frac{1}{\tau\sqrt{N}}\}}\right) \geq \frac{1}{\log(\lambda_1)-\log(\lambda_2)}\log\left(\frac{\tan\theta_1^{(0)}|\lambda_1|}{(\lambda_2-\lambda_3)\min\{\rho,\frac{1}{\tau\sqrt{N}}\}}\right).$$

Moreover, $\max\{\lambda^2|\lambda_2|\phi^L, N|\lambda_1^3|\phi^L\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1}\}$ should less than $(\lambda_2-\lambda_3)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}$, which means

$$L = \Omega\left(\frac{1}{\log\left(\phi^{-1}\right)}\log\left(\frac{N|\lambda_1^3|\left|\frac{\lambda_1}{\lambda_N}\right|^{T_1}}{(\lambda_2-\lambda_3)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}}\right)\right). \tag{38}$$

Using the argument in **Theorem 1**, the number of power iteration for estimating $(\lambda_2, \boldsymbol{u}_2)$ is

$$T_2 = \Omega\left(\frac{\lambda_2}{\lambda_2-\lambda_3}\log\left(\frac{\tau\sqrt{N}}{\rho}\right)\right).$$

Substitute T_1 in to Equation (38) and we can obtain

$$\begin{aligned}
L &= \Omega\left(\frac{1}{\log\left(\phi^{-1}\right)}\log\left(\frac{N|\lambda_1^3|\left(\frac{\lambda_1}{\lambda_N}\right)^{T_1}}{(\lambda_2-\lambda_3)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}}\right)\right) \\
&= \Omega\left(\frac{1}{\log\left(\phi^{-1}\right)}\left(T_1\log\left(\left|\frac{\lambda_1}{\lambda_N}\right|\right)+\log\left(\frac{N|\lambda_1^3|}{(\lambda_2-\lambda_3)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}}\right)\right)\right), \\
&= \Omega\left(\frac{1}{\log\left(\phi^{-1}\right)}\left(\frac{1}{\log\left(\frac{\lambda_1}{\lambda_2}\right)}\log\left(\frac{\tau\sqrt{N}|\lambda_1|}{(\lambda_2-\lambda_3)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}}\right)\log\left(\left|\frac{\lambda_1}{\lambda_N}\right|\right)+\log\left(\frac{N|\lambda_1^3|}{(\lambda_2-\lambda_3)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}}\right)\right)\right) \\
&= \Omega\left(\frac{1}{\log\left(\phi^{-1}\right)}\frac{\log\left(\left|\frac{\lambda_1}{\lambda_N}\right|\right)}{\log\left(\frac{\lambda_1}{\lambda_2}\right)}\log\left(\frac{\tau N|\lambda_1^3|}{(\lambda_2-\lambda_3)\min\{\rho, \frac{1}{\tau\sqrt{N}}\}}\right)\right)
\end{aligned}$$

□

3 Auxiliary Lemmas: Stochastic Block Model

Consider a symmetric matrix $\boldsymbol{P} \in \mathbb{R}^{N\times N}$ that follows a stochastic block model (SBM) denoted as SBM$(N, K, p = \frac{a\log N}{N}, q = \frac{b\log N}{N})$. Each entry P_{ij} of $\boldsymbol{P}$ is defined as follows

$$P_{ij} = \begin{cases} p, & \text{if } y_i^\star = y_j^\star, \\ q, & \text{if } y_i^\star \neq y_j^\star. \end{cases},$$

where $y_i^\star$ and $y_j^\star$ represent the block assignments of nodes i and j, respectively. It is important to note that the matrix $\boldsymbol{P}$ differs from $\mathbb{E}[\boldsymbol{A}]$ due to the diagonal terms.

Let $\lambda_1^\star \geq \ldots \geq \lambda_K^\star$ denote the eigenvalues of $\boldsymbol{P}$ and $\boldsymbol{u}_1^\star, \ldots, \boldsymbol{u}_K^\star \in \mathbb{R}^N$ denote the corresponding eigenvectors. We define $\boldsymbol{U}^\star = [\boldsymbol{u}_1^\star, \ldots, \boldsymbol{u}_K^\star] \in \mathbb{R}^{N\times K}$ as the leading eigenspace and $\boldsymbol{\Lambda}^\star = \operatorname{diag}(\lambda_1^\star, \ldots, \lambda_K^\star) \in \mathbb{R}^{K\times K}$ as the diagonal matrix containing the leading eigenvalues. Therefore, the matrix $\boldsymbol{U}^\star\boldsymbol{\Lambda}^\star$ satisfies the following property:

Lemma 3.1. *([6, Lemma 1]) Let the notation be as above and the matrix $\boldsymbol{U}^\star\boldsymbol{\Lambda}^\star \in \mathbb{R}^{N\times K}$ has K unique rows. To be more specific, there exist $\boldsymbol{\mu}_1^\star, \ldots, \boldsymbol{\mu}_K^\star \in \mathbb{R}^{1\times K}$ such that $(\boldsymbol{U}^\star\boldsymbol{\Lambda}^\star)_i = \boldsymbol{\mu}_{y_i^\star}^\star$ for all $i \in [N]$. In addition,*

$$\|\boldsymbol{\mu}_a^\star - \boldsymbol{\mu}_b^\star\| = \sqrt{n_a+n_b}\,(p-q),$$

and $\|\boldsymbol{\mu}_a^\star\|_2^2 = \left(p^2-q^2\right)n_a + q^2N$, for all $a, b \in [K]$ such that $a \neq b$.

Although the adjacency matrix $\boldsymbol{A} \in \mathbb{R}^{N\times N}$ and the corresponding eigenspace $\boldsymbol{U}, \boldsymbol{\Lambda}$ differ from $\boldsymbol{P}$ and $\boldsymbol{U}^\star, \boldsymbol{\Lambda}^\star$, the deviation from the expected matrices and the ideal eigenspace are bounded. In the following sections, we introduce several important Lemmas.

Lemma 3.2. *([7, Theorem 5.2]) Let $\boldsymbol{A} \in \mathbb{R}^{N\times N}$ be the adjacency matrix of a random graph on N nodes whose edges are sampled independently. Let $\mathbb{E}[\boldsymbol{A}] = (p_{ij})_{i,j=1,2,\cdots,N}$ and assume that $N \max_{ij} p_{ij} \leq d$ for $d \geq c_0 \log N$ and $c_0 > 0$. Then, for any $r > 0$ there exists a constant $C = C(r, c_0)$ such that*

$$\|\boldsymbol{A} - \mathbb{E}[\boldsymbol{A}]\|_2 \leq C\sqrt{d}$$

with probability at least $1 - N^{-r}$.

Lemma 3.2 provides a spectral bound for binary symmetric random matrices on the deviation from the expectation. Since we consider SBM with parameters $p, q = \Omega(\dfrac{\log N}{N})$ with no self-loop, we have the following lemma with $d = \Omega(\log N)$.

Lemma 3.3. *Assume $\boldsymbol{A} \sim \mathrm{SBM}(N, K, p = \frac{a\log N}{N}, q = \frac{b\log N}{N})$ and thus $\boldsymbol{P} \in \mathbb{R}^{N\times N}$, where $P_{ij} = p\mathbb{I}\{y_i^\star = y_j^\star\} + q\mathbb{I}\{y_i^\star \neq y_j^\star\}$, $\forall i, j \in [N]$. There exists a constant $C_3 > 0$ such that*

$$\|\boldsymbol{A} - \boldsymbol{P}\|_2 \leq C_3\sqrt{\log N}$$

holds with probability at least $1 - N^{-3}$.

Proof. Since $\mathbb{E}[\boldsymbol{A}]_{ij} = \boldsymbol{P}_{ij}, \forall i \neq j$ and $\mathbb{E}[\boldsymbol{A}]_{ii} = 0, \forall i \in [N]$.

$$\begin{aligned}
\|\boldsymbol{A} - \boldsymbol{P}\|_2 &= \|\boldsymbol{A} - \mathbb{E}[\boldsymbol{A}] + \mathbb{E}[\boldsymbol{A}] - \boldsymbol{P}\|_2 \\
&\leq \|\boldsymbol{A} - \mathbb{E}[\boldsymbol{A}]\|_2 + \|\mathbb{E}[\boldsymbol{A}] - \boldsymbol{P}\|_2 \\
&\leq C\sqrt{\log N} + \frac{a\log N}{N} \\
&= \mathcal{O}\left(\sqrt{\log N}\right).
\end{aligned}$$

The inequality holds due to **Lemma 3.2** and $\sqrt{\log N} > \frac{\log N}{N}$. □

To show the relationship between $\boldsymbol{A}$ and $\boldsymbol{P}$ in terms of the eigenvalues, we utilize **Lemma 3.4**, **Lemma 3.5** and **Lemma 3.6**.

Lemma 3.4 (Theorem of perturbations of the spectrum)**.** *If $\|\boldsymbol{A} - \boldsymbol{P}\|_2 \leq \epsilon$, for some $\epsilon > 0$. Then it follows that all the eigenvalues of $\boldsymbol{A} - \boldsymbol{P}$ are bounded in absolute value by ϵ. Applying Weyl's inequality [8], it follows that the spectra of the symmetric matrices $\boldsymbol{A}$ and $\boldsymbol{P}$ are close in the sense that:*

$$|\lambda_i(\boldsymbol{A}) - \lambda_i(\boldsymbol{P})| \leq \epsilon, \ \forall i \in [N].$$

Define $\boldsymbol{Y}^\star \in \{0,1\}^{N\times K}$ to be a matrix such that $Y_{ij}^\star = \mathbb{I}\{y_i^\star = y_j^\star\}$. In addition, $\boldsymbol{B} \in \mathbb{R}^{K\times K}$ such that $B_{cd} = p\mathbb{I}\{c = d\} + q\mathbb{I}\{c \neq d\}$ for all $c, d \in [K]$. Then we can verify

$$\boldsymbol{P} = \boldsymbol{Y}^\star \boldsymbol{B} (\boldsymbol{Y}^\star)^{\mathtt{H}},$$

where the rank of $\boldsymbol{P}$ is K and $\lambda_j(\boldsymbol{P}) = 0, \forall j \geq K + 1$. In the following, we consider $\boldsymbol{A} \sim \mathrm{SBM}(N, K, p = \frac{a\log N}{N}, q = \frac{b\log N}{N})$, and thus $\boldsymbol{P} \in \mathbb{R}^{N\times N}$, where

$$P_{ij} = p\mathbb{I}\{y_i^\star = y_j^\star\} + q\mathbb{I}\{y_i^\star \neq y_j^\star\}, \ \forall i, j \in [N].$$

Then, we derive the upper bound for eigenvalues of $\boldsymbol{P}$.

Lemma 3.5 (Upper bound for eigenvalues)**.** *Let the notation be as above. Then all eigenvalues of $\boldsymbol{P}$ have the following property:*

$$\forall i = 1, 2, ...N, \lambda_i(\boldsymbol{P}) \leq \mathcal{O}(\log N)$$

Proof. According to Gershgorin circle theorem, the largest eigenvalue is upper bounded by the largest absolute row sum or column sum. By the definition of $\boldsymbol{P}$, the largest absolute row sum or column sum is

$$\max_i \sum_j P_{ij} = \max_j \sum_i P_{ji} = C \log N,$$

where C is a constant. $\square$

Next, we rewrite Lemma 3 and Lemma 5 in [6] to bound the eigenvalues of the adjacency matrix $\boldsymbol{A}$ using $\boldsymbol{P}$.

Lemma 3.6 (Restate Lemma 3 and 5 [6]). *Assume $\boldsymbol{A} \sim \mathrm{SBM}(N, K, p = \frac{a \log N}{N}, q = \frac{b \log N}{N})$, and thus $\boldsymbol{P} \in \mathbb{R}^{N \times N}$, where $P_{ij} = p\mathbb{I}\{y_i^\star = y_j^\star\} + q\mathbb{I}\{y_i^\star \neq y_j^\star\}$, $\forall i, j \in [N]$. We have $\forall i = 1, 2, ...K, \lambda_i(\boldsymbol{P}) = \Theta(\log N)$ and with probability at least $1 - N^{-3}$*

$$\max_{i=1,2,\ldots K} |\lambda_i(\boldsymbol{A}) - \lambda_i(\boldsymbol{P})| \leq \|\boldsymbol{A} - \boldsymbol{P}\|_2 \leq \mathcal{O}(\sqrt{\log N})$$

$$\max_{i \geq K+1} |\lambda_i(\boldsymbol{A})| \leq \|\boldsymbol{A} - \boldsymbol{P}\|_2 \leq \mathcal{O}(\sqrt{\log N}).$$

Then

$$\min_{i=1,2,\ldots K} \lambda_i(\boldsymbol{A}) = \Theta(\log N), \min_{i=1,2,\ldots K} |\lambda_i(\boldsymbol{A})| \quad > \max_{j \geq K+1} |\lambda_j(\boldsymbol{A})|.$$

Proof. Define $\boldsymbol{\Delta} = \mathrm{diag}\left(\sqrt{n_1}, \ldots, \sqrt{n_K}\right) \in \mathbb{R}^{K \times K}$. Note that

$$\boldsymbol{P} = \boldsymbol{Y}^\star \boldsymbol{B} (\boldsymbol{Y}^\star)^{\mathtt{H}} = \boldsymbol{Y}^\star \boldsymbol{\Delta}^{-1} \left(\boldsymbol{\Delta B \Delta}\right) \boldsymbol{\Delta}^{-1} (\boldsymbol{Y}^\star)^{\mathtt{H}}, \text{ and } \boldsymbol{Y}^\star \boldsymbol{\Delta}^{-1} \in \mathcal{O}(N, K),$$

where $\mathcal{O}(N, K)$ is the set of all $N \times K$ matrices with orthogonal columns. Apply SVD to the matrix $\boldsymbol{\Delta B \Delta}$ and we obatin $\boldsymbol{\Delta B \Delta} = \boldsymbol{W} \check{\boldsymbol{\Lambda}} \boldsymbol{W}^{\mathtt{H}}$ for some $\boldsymbol{W} \in \mathcal{O}(K, K)$ and some diagonal matrix $\check{\boldsymbol{\Lambda}}$. Similarly, applying SVD to the matrix $\boldsymbol{P}$ and we obtain $\boldsymbol{P} = \boldsymbol{W}' \boldsymbol{\Lambda}^\star \boldsymbol{W}'$ for some $\boldsymbol{W}' \in \mathcal{O}(K, K)$ and some diagonal matrix $\boldsymbol{\Lambda}^\star$. Since $\boldsymbol{Y}^\star \boldsymbol{\Delta}^{-1} \boldsymbol{W} \in \mathcal{O}(N, K)$ and

$$\boldsymbol{P} = \boldsymbol{W}' \boldsymbol{\Lambda}^\star \boldsymbol{W}' = \left(\boldsymbol{Y}^\star \boldsymbol{\Delta}^{-1} \boldsymbol{W}\right) \check{\boldsymbol{\Lambda}} \left(\boldsymbol{Y}^\star \boldsymbol{\Delta}^{-1} \boldsymbol{W}\right)^{\mathtt{H}},$$

we have $\boldsymbol{\Lambda}^\star = \check{\boldsymbol{\Lambda}} = \boldsymbol{W}^{\mathtt{H}} \boldsymbol{\Delta B \Delta W}$. In the other words, $\lambda_1(\boldsymbol{P}), ..., \lambda_K(\boldsymbol{P})$ are the eigenvalues of $\boldsymbol{\Lambda B \Lambda}$. Then

$$\begin{aligned}
\lambda_K(\boldsymbol{P}) &\geq \min_{\boldsymbol{v} \in \mathbb{R}^K, \|\boldsymbol{v}\|_2 = 1} \boldsymbol{v}^{\mathtt{H}} \boldsymbol{\Delta B \Lambda v} = \min_{\boldsymbol{v} \in \mathbb{R}^K, \|\boldsymbol{v}\|_2 = 1} (\boldsymbol{v\Delta})^{\mathtt{H}} \boldsymbol{B} (\boldsymbol{\Delta v}) \\
&\geq \left(\min_{f \in [K]} n_f\right) \min_{\boldsymbol{v} \in \mathbb{R}^K, \|\boldsymbol{v}\|_2 = 1} \boldsymbol{v}^{\mathtt{H}} \boldsymbol{B v} \\
&= \left(\min_{f \in [K]} n_f\right) \min_{\boldsymbol{v} \in \mathbb{R}^K, \|\boldsymbol{v}\|_2 = 1} \left(q \left(\boldsymbol{v}^{\mathtt{H}} \mathbf{1}_K\right)^2 + (p - q) \|\boldsymbol{v}\|_2^2\right) \\
&\geq (p - q) \left(\min_{f \in [K]} n_f\right) = \Omega(\log N).
\end{aligned}$$

Observing **Lemma 3.5**, the upper bound and the lower bound of first K eigenvalues are in the same order, which implies

$$\min_{i=1,2,\ldots K} \lambda_i(\boldsymbol{A}) = \Theta(\log N).$$

On the other hand, **Lemma 3.3** and **Lemma 3.4** show that $\max_{i=1,2,\ldots K} |\lambda_i(\boldsymbol{A}) - \lambda_i(\boldsymbol{P})|$ and $\max_{i \geq K+1} |\lambda_i(\boldsymbol{A})|$ are upper bounded by $\|\boldsymbol{A} - \boldsymbol{P}\|_2 = \mathcal{O}\left(\sqrt{\log N}\right)$ with probability at least $1 - N^{-3}$. Since $\forall i \geq K+1, \ \lambda_i(\boldsymbol{P}) = 0$, we can derive

$$\min_{i=1,2,\ldots K} |\lambda_i(\boldsymbol{A})| = \Theta(\log N) > \max_{j \geq K+1} |\lambda_j(\boldsymbol{A})| = \mathcal{O}(\sqrt{\log N}).$$

$\square$

Remark 3.1. *For a symmetric real matrix, singular values equal the absolute value of eigenvalues. Moreover, Lemma 3.6 shows $\lambda_K(\boldsymbol{P}) = \Omega(\log N)$, which means*

$$\sigma_{\min}(\boldsymbol{P}) = \lambda_K(\boldsymbol{P}) = \Omega(\log N).$$

4 Auxiliary Lemmas: Power method

We consider a symmetric graph matrix $\boldsymbol{Z} \in \mathbb{R}^{N\times N}$. Now, we present well-known theoretical results for the power method. The centralized power method performs the following updates in each iteration

$$\begin{aligned}
\hat{\boldsymbol{u}}_1^{(t+1)} &= \boldsymbol{Z}\boldsymbol{q}_1^{(t)} \\
\boldsymbol{q}_1^{(t+1)} &= \hat{\boldsymbol{u}}_1^{(t+1)} / \,\|\hat{\boldsymbol{u}}_1^{(t+1)}\|_2 \\
\gamma_1^{(t+1)} &= (\boldsymbol{q}_1^{(t+1)})^{\mathrm{H}}\boldsymbol{Z}\boldsymbol{q}_1^{(t+1)}
\end{aligned} \tag{39}$$

To initialize $\boldsymbol{q}_1^{(0)} \in \mathbb{R}$ in a distributed manner, each node i selects a random value $q_{1,(i)}^0$ from a Gaussian distribution with zero mean and $\frac{1}{N}$ variance. The following Lemma provides a proof that the length of the initial vector is bounded by $\frac{3}{2}$ with a high probability.

Lemma 4.1 (Length of random gaussian vectors are bounded)**.** *Let $\boldsymbol{x} \in \mathbb{R}^N$ be a vector such that $\forall i \in [N], x_i \sim \mathcal{N}\left(0, \frac{1}{N}\right)$. Then, with probability at least $1 - e^{-\left(\frac{5}{8} - \log\frac{3}{2}\right)N}$, we have $\|\boldsymbol{x}\|_2 < \frac{3}{2}$.*

Proof. To prove that $\|\boldsymbol{x}\|_2 \leq \frac{3}{2}$, we utilize the Chernoff bound, which states:

$$\mathbb{P}\left(\sum_{i=1}^{N} x_i^2 - \mathbb{E}[x_i^2] \geq a\right) \leq \inf_{s\geq 0} e^{-sa}\mathbb{E}[e^{s(x_1^2 - \frac{1}{N})}]^N.$$

We observe that the moment-generating function is given by:

$$\begin{aligned}
\mathbb{E}[e^{s(x_i^2-\frac{1}{N})}] &= \frac{\sqrt{N}}{\sqrt{2\pi}}e^{-\frac{s}{N}}\int_{-\infty}^{\infty} e^{sx_1^2 - \frac{N}{2}x_1^2}dx \\
&= \frac{\sqrt{N}}{\sqrt{2\pi}}\frac{\sqrt{2}}{\sqrt{N-2s}}e^{-\frac{s}{N}}\int_{-\infty}^{\infty} e^{-w^2}dw \\
&= \frac{1}{\pi}\left(\frac{N}{N-2s}\right)^{1/2} e^{-\frac{s}{N}}\sqrt{\pi} \\
&= \left(\frac{N}{N-2s}\right)^{1/2} e^{-\frac{s}{N}}.
\end{aligned}$$

Therefore, we can write

$$\mathbb{P}\left(\sum_{i=1}^{N} x_i^2 - \mathbb{E}[x_i^2] \geq a\right) \leq \inf_{s\geq 0} e^{-sa}\left(\frac{N}{N-2s}\right)^{N/2} e^{-s}.$$

Due to $\sum_{i=1}^N \mathbb{E}[x_i^2] = 1$, we set $a = \frac{5}{4}$ such that $\sum_{i=1}^N x_i^2 \geq \frac{9}{4}$. By performing differential, we obtain the optimal $s = \frac{5}{18}N$ and the corresponding upper bound:

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{N} x_i^2 \geq \frac{9}{4}\right) &= \mathbb{P}\left(\|\boldsymbol{x}\|_2^2 \geq \frac{9}{4}\right) \\
&\leq \exp\left(-\frac{5}{8}N + N\log\frac{3}{2}\right),
\end{aligned}$$

which implies that $\|\boldsymbol{x}\|_2 < \frac{3}{2}$ with probability at least $1 - e^{-\left(\frac{5}{8} - \log\frac{3}{2}\right)N}$. ∎

To analyze the convergence of the power method, we introduce Theorem 8.2.1 from [9] as **Lemma 4.2**.

Lemma 4.2. *([9, Theorem 8.2.1]) Suppose the graph matrix $\boldsymbol{Z} \in \mathbb{R}^{N\times N}$ is symmetric and that*

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{Z}\boldsymbol{U} = \mathrm{diag}\left(\lambda_1, \ldots, \lambda_N\right)$$

where $\boldsymbol{U} = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_N] \in \mathbb{R}^{N \times N}$ *is orthogonal and* $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_N|$*. Let the vector* $\boldsymbol{q}_1^{(t)} = \frac{\boldsymbol{Z}\boldsymbol{q}_1^{(t-1)}}{\|\boldsymbol{Z}\boldsymbol{q}_1^{(t-1)}\|_2}$ *and define* $\theta_1^{(t)} \in [0, \pi/2]$ *by* $\left(\cos\theta_1^{(t)}\right) = \left|\boldsymbol{u}_1^{\mathtt{H}}\boldsymbol{q}_1^{(t)}\right|$ *If* $\left(\cos\theta_1^{(0)}\right) \neq 0$*, then for* $t = 0, 1, \ldots$ *we have*

$$\left|\sin\theta_1^{(t)}\right| \leq \left(\tan\theta_1^{(0)}\right)\left|\frac{\lambda_2}{\lambda_1}\right|^t$$

This implies estimated eigenvector $\boldsymbol{q}_1^{(t)}$ will converge to true eigenvector $\boldsymbol{u}_1$ as t is sufficiently large. Recall that we have adjusted the update rules of the vanilla power method to create the lightweight power method, which performs the following updates at each iteration:

$$\begin{aligned}
\hat{\boldsymbol{u}}_1^{(t+1)} &= \boldsymbol{Z}\boldsymbol{q}_1^{(t)} \\
\gamma_1^{(t+1)} &= \|\hat{\boldsymbol{u}}_1^{(t+1)}\|_2 \\
\boldsymbol{q}_1^{(t+1)} &= \hat{\boldsymbol{u}}_1^{(t+1)} / \|\hat{\boldsymbol{u}}_1^{(t+1)}\|_2
\end{aligned} \tag{40}$$

Lemma 4.3. *Suppose the graph matrix* $\boldsymbol{Z} \in \mathbb{R}^{N \times N}$ *is symmetric and that*

$$\boldsymbol{U}^{\mathtt{H}}\boldsymbol{Z}\boldsymbol{U} = \operatorname{diag}(\lambda_1, \ldots, \lambda_N)$$

where $\boldsymbol{U} = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_N] \in \mathbb{R}^{N \times N}$ *is orthogonal and* $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_N|$*. Let the vector* $\boldsymbol{q}_1^{(t)} = \frac{\boldsymbol{Z}\boldsymbol{q}_1^{(t-1)}}{\|\boldsymbol{Z}\boldsymbol{q}_1^{(t-1)}\|_2}$*, the scalar* $\gamma_1^{(t)} = \|\boldsymbol{Z}\boldsymbol{q}_1^{(t-1)}\|_2$ *and define* $\left(\cos\theta_1^{(t)}\right) = \boldsymbol{u}_1^{\mathtt{H}}\boldsymbol{q}_1^{(t)}$*. Assume* $\cos\theta_1^{(0)} \geq 0$ *and* $\operatorname{sgn}(\lambda_i) = \operatorname{sgn}(\gamma_1^{(t)})$*, then for* $t = 0, 1, \ldots$ *we have*

$$\begin{aligned}
\|\boldsymbol{u}_1 - \boldsymbol{q}_1^{(t)}\|_2 &\leq \sqrt{2}\left(\tan\theta_1^{(0)}\right)\left|\frac{\lambda_2}{\lambda_1}\right|^t \\
|\lambda_1 - \gamma_1^{(t)}| &\leq \left|\frac{\lambda_1}{\lambda_N}\right| \max_{m=2,\ldots,N} |(\lambda_1 - \lambda_m)| \left(\tan\theta_1^{(0)}\right)^2 \left|\frac{\lambda_2}{\lambda_1}\right|^{2t-2}
\end{aligned}$$

Proof. We first prove the convergence of the eigenvector. Note that $\boldsymbol{u}_1^{(t)}$ and $\boldsymbol{u}_1$ are unit vectors, so we have:

$$\|\boldsymbol{u}_1 - \boldsymbol{q}_1^{(t)}\|_2^2 = 2\left(1 - \boldsymbol{u}_1^{\mathtt{H}}\boldsymbol{q}_1^{(t)}\right) \leq 2\left(1 - \left(\boldsymbol{u}_1^{\mathtt{H}}\boldsymbol{q}_1^{(t)}\right)^2\right) = 2\left(\sin\theta_1^{(t)}\right)^2.$$

Since **Lemma 4.2** guarantees the convergence of $\sin\theta_1^{(t)}$, we obtain:

$$\|\boldsymbol{u}_1 - \boldsymbol{q}_1^{(t)}\|_2 \leq \sqrt{2}\left|\sin\theta_1^{(t)}\right| \leq \sqrt{2}\left(\tan\theta_1^{(0)}\right)\left|\frac{\lambda_2}{\lambda_1}\right|^t.$$

Now we show the estimated error of eigenvalues with exponential decay. Since we have assumed $\operatorname{sgn}(\lambda_1) = \operatorname{sgn}(\gamma_1^{(t)})$, by the definition,

$$\begin{aligned}
|\gamma_1^{(t)}| &= \|\boldsymbol{Z}\boldsymbol{u}_1^{(t-1)}\|_2 = \sqrt{(\boldsymbol{u}_1^{(t-1)})^H\boldsymbol{Z}^H\boldsymbol{Z}\boldsymbol{u}_1^{(t-1)}} \\
&= \left(\frac{(\boldsymbol{u}_1^{(0)})^H\boldsymbol{Z}^{2t}\boldsymbol{u}_1^{(0)}}{(\boldsymbol{u}_1^{(0)})^H\boldsymbol{Z}^{2t-2}\boldsymbol{u}_1^{(0)}}\right)^{0.5}
\end{aligned} \tag{41}$$

Since the initial vector $\boldsymbol{u}_1^{(0)}$ can be written as $\sum_{m=1}^{N} \alpha_m \boldsymbol{u}_m$. Then, we have

$$\begin{aligned}
|\gamma_1^{(t)}| &= \sqrt{\frac{(\boldsymbol{u}_1^{(0)})^H\boldsymbol{Z}^{2t}\boldsymbol{u}_1^{(0)}}{(\boldsymbol{u}_1^{(0)})^H\boldsymbol{Z}^{2t-2}\boldsymbol{u}_1^{(0)}}} = \sqrt{\frac{\sum_{m=1}^{N}\alpha_m^2\lambda_m^{2t}}{\sum_{m=1}^{N}\alpha_m^2\lambda_m^{2t-2}}} \\
&\geq \sqrt{\lambda_1^2\frac{\sum_{m=1}^{N}\alpha_m^2\lambda_m^{2t-2}}{\sum_{m=1}^{N}\alpha_m^2\lambda_m^{2t-2}}} = |\lambda_N|\sqrt{\frac{\sum_{m=1}^{N}\alpha_m^2\lambda_m^{2t-2}}{\sum_{m=1}^{N}\alpha_m^2\lambda_m^{2t-2}}} = |\lambda_N|.
\end{aligned} \tag{42}$$

On the other hand, we have

$$
\begin{aligned}
|\gamma_1^{(t)}| &= \sqrt{\frac{(\boldsymbol{u}_1^{(0)})^H \boldsymbol{Z}^{2t} \boldsymbol{u}_1^{(0)}}{(\boldsymbol{u}_1^{(0)})^H \boldsymbol{Z}^{2t-2} \boldsymbol{u}_1^{(0)}}} = \sqrt{\frac{\sum_{m=1}^{N} \alpha_m^2 \lambda_m^{2t}}{\sum_{m=1}^{N} \alpha_m^2 \lambda_m^{2t-2}}} \\
&\leq \sqrt{\lambda_1^2 \frac{\sum_{m=1}^{N} \alpha_m^2 \lambda_m^{2t-2}}{\sum_{m=1}^{N} \alpha_m^2 \lambda_m^{2t-2}}} = |\lambda_1| \sqrt{\frac{\sum_{m=1}^{N} \alpha_m^2 \lambda_m^{2t-2}}{\sum_{m=1}^{N} \alpha_m^2 \lambda_m^{2t-2}}} = |\lambda_1|.
\end{aligned} \tag{43}
$$

Therefore, we have

$$
|\lambda_1| \geq |\gamma_1^{(t)}| \geq |\lambda_N|.
$$

Moreover, we can derive the upper bound as follows:

$$
\begin{aligned}
(\lambda_1)^2 - (\gamma_1^{(t)})^2 &= \frac{\lambda_1^2 \sum_{m=1}^{N} \alpha_m^2 \lambda_1^{2t-2} - \sum_{m=1}^{N} \alpha_m^2 \lambda_m^{2t}}{\sum_{m=1}^{N} \alpha_m^2 \lambda_m^{2t-2}} = \frac{\sum_{m=1}^{N} \alpha_m^2 \lambda_m^{2t-2} (\lambda_1^2 - \lambda_m^2)}{\sum_{m=1}^{N} \alpha_m^2 \lambda_m^{2t-2}} \\
&\leq \max_{m=2,\ldots,N} (\lambda_1^2 - \lambda_m^2) \frac{\sum_{m=2}^{N} \alpha_m^2 \lambda_m^{2t-2}}{\alpha_1^2 \lambda_1^{2t-2}} \\
&\leq \max_{m=2,\ldots,N} (\lambda_1^2 - \lambda_m^2) \frac{1}{\alpha_1^2} \sum_{m=2}^{N} \alpha_m^2 \left(\frac{\lambda_m}{\lambda_1} \right)^{2t-2} \\
&\leq \max_{m=2,\ldots,N} (\lambda_1^2 - \lambda_m^2) \frac{1 - \alpha_1^2}{\alpha_1^2} \left| \frac{\lambda_2}{\lambda_1} \right|^{2t-2} \\
&\leq \max_{m=2,\ldots,N} |\lambda_1 + \lambda_m| |\lambda_1 - \lambda_m| \left(\tan \theta_1^{(0)} \right)^2 \left| \frac{\lambda_2}{\lambda_1} \right|^{2t-2} \\
&\leq 2|\lambda_1| \max_{m=2,\ldots,N} |\lambda_1 - \lambda_m| \left(\tan \theta_1^{(0)} \right)^2 \left| \frac{\lambda_2}{\lambda_1} \right|^{2t-2}
\end{aligned} \tag{44}
$$

Note that $(\lambda_1)^2 - (\gamma_1^{(t)})^2 = (\lambda_1 + \gamma_1^{(t)})(\lambda_1 - \gamma_1^{(t)}) \geq 2|\lambda_N||\lambda_1 - \gamma_1^{(t)}|$. Combining with Equation (44), we have

$$
|\lambda_1 - \gamma_1^{(t)}| \leq \left| \frac{\lambda_1}{\lambda_N} \right| \max_{m=2,\ldots,N} |\lambda_1 - \lambda_m| \left(\tan \theta_1^{(0)} \right)^2 \left| \frac{\lambda_2}{\lambda_1} \right|^{2t-2} \tag{45}
$$

□

References

[1] S. Boyd, P. Diaconis, and L. Xiao, "Fastest mixing markov chain on a graph," *SIAM review*, vol. 46, no. 4, pp. 667–689, 2004.

[2] T. Tao and V. Vu, "Random matrices have simple spectrum," *Combinatorica*, vol. 37, pp. 539–553, 2017.

[3] M. Hardt and E. Price, "The noisy power method: A meta algorithm with applications," *Advances in neural information processing systems*, vol. 27, 2014.

[4] M. Rudelson and R. Vershynin, "Smallest singular value of a random rectangular matrix," *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, vol. 62, no. 12, pp. 1707–1739, 2009.

[5] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE transactions on information theory*, vol. 52, no. 6, pp. 2508–2530, 2006.

[6] A. Y. Zhang, "Fundamental limits of spectral clustering in stochastic block models," *arXiv preprint arXiv:2301.09289*, 2023.

[7] J. Lei and A. Rinaldo, "Consistency of spectral clustering in stochastic block models," *The Annals of Statistics*, pp. 215–237, 2015.

[8] H. Weyl, "Das asymptotische verteilungsgesetz der eigenwerte linearer partieller differentialgleichungen (mit einer anwendung auf die theorie der hohlraumstrahlung)," *Mathematische Annalen*, vol. 71, no. 4, pp. 441–479, 1912.

[9] G. H. Golub and C. F. Van Loan, *Matrix computations.* JHU press, 2013.