CONSENSUS AND BIAS

MURI L.A. GROUP

Abstract. In this paper, we study consensus formation in a gossip-based opinion dynamics model. We consider whether social biases towards certain information can affect the speed of convergence to, and the nature of, the consensus reached. In particular, we focus on the case where individuals are "information seeking," meaning that individuals communicate more often with individuals who have an opinion that is different than their own. In this case, we find that the convergence to consensus under the information seeking bias is [[hopefully sublinear but ??]] faster for [[certain graphs... not sure yet]] than convergence when individuals communicate at a constant rate in time. One implication of these findings is that consensus-formation in social networks can be greatly accelerated if individuals have biases based on others' opinions that cause them to communicate more frequently with others. [[Insert something like "in order to explain slow or fast consensus among individuals it is important to incorporate biases that may affect communication rates.]]

1. Introduction

[[¶1 Opinion formation in social networks. This paragraph needs work to make tight.]] The formation of individual opinion in a given setting is a complicated process involving the interplay between many factors including 1.) the structural and social characteristics of the setting, 2.) the structure of an individual's social network, 3.) the frequency with which an individual communicates with its social network, and 4.) initial opinions [[cite sociology literature]]. Opinion formation itself is the combination of at least two separate components, including 1.) the time scale at which an individual's opinion changes and 2.) the opinion held by an individual at a certain point in time [[sociology literature]]. The study of opinion formation is characterized by understanding how the stated factors affect the two components of opinion formation.

[[**¶2-¶**k mini lit review the aspects of opinion formation that have been explored in the past.]] In particular, it is known that certain assumptions on the opinion formation process result in an eventual convergence to a consensus. [[What settings have been considered? What choices of 2-5 above have been explored? What affects do these choices have on the "components" 1-2? Structure paragraphs based on the numbered points above.]] [[Make the point that an opinion dynamics model is a combination of 1.) a rule for how an individual changes their opinion when an update occurs and 2.) a rule for which nodes update at a given point in time]]

The concern of this document is how the factors 2.) and 3.) above, i.e. the structure of an individual's social network and the frequency of communication, affects the component 1.) above, the time scale of opinion formation, under the following assumptions:

- **A1** Opinions are represented by real numbers, and the distance between two groups of opinions is given by the Euclidean distance.
- **A2** Communication occurs between pairs of individuals.
- **A3** When a pair communicates, they come to a consensus, meaning that the pair of nodes has the same opinion.
- **A4** The resulting new opinions of the individuals in the social network, including the particular consensus agreed upon by the communicating pair of individuals, after communication between a pair of nodes

Date: March 2019.

minimizes the distance between the old opinions and any new opinion. In other words, the individuals in the network, including the communicating pair, are assumed to change their opinions as little as possible when a pair of nodes comes to a consensus.

[[What is the sociological precedent for these assumptions? 1. is a property of the "setting" in which individuals update (online social network, in person, etc.), 2. is a property of the opinions and the distance models how individuals perceive how different opinions are from one another, 3. assumes that communication between individuals results in a consensus between the individuals (in what situations might this hold?), and 4. Assumes that individuals wish to change their opinions as little as possible.]]

As will be shown later, the assumptions above can be used to derive how the opinions of the nodes in the network change upon an update by a pair of nodes. To complete the model, one must specify the time-dynamics: 1.) when do updates occur and 2.) which nodes communicate when an update occurs. The assumptions made in this document have the effect that results in terms of the discrete number of updates can be converted to results about continuous time; in other words, the results depend only on the events and their ordering rather than the time intervals between events. Accordingly, specification of 2.) is all that is required.

The opinion dynamics model, consisting of an update rule derived using the assumptions **A1**-**A4** and a rule for choosing which nodes communicate, converges to an even consensus, meaning that the opinions of the nodes in the network converge to the same value and that the value is the average of the initial opinions of the nodes, under mild assumptions on which pairs of nodes communicate and on the network. In this document, we investigate the time-scale of convergence to consensus under the assumption that the nodes are "information-seeking," meaning that the nodes communicate more often with nodes that have an opinion that is more different than their own opinion. The particular communication model we use to investigate the affect of this assumption completes the opinion dynamics model by determining the pair of nodes that communicate at an update time.

[[TODO: Summary of results in this document about information seeking-behavior.]]

2. Model and General Results

2.1. **Derivation of Update Rule.**

Let G be a connected undirected graph with vertex set $\mathcal{V}(G)$ of size n and edge set $\mathcal{E}(G)$ of size m. We assume that G will be sufficent as a model of the social network we study. Denote by $x_i(t)$ the opinion of node i at time t and denote the vector of all $x_i(t)$ for the nodes in the graph by $x(t)$. As stated in **A1**, we assume that $x_i(t) \in \mathbb{R}$. We assume that update events are generated by a Poisson process. As stated in the appendix [[reference here]], this assumption allows us to develop results for discrete time opinion dyanmics, and then later adapt these discrete time results to the continuous time setting. Accordingly, assume that a countable sequence of update times $0 = t_0 < t_1 < \cdots$ is given and denote $x_i(k) = x_i(t^k)$. Denote by $M \in \{-1, 0, 1\}^{2m \times n}$ the incidence matrix, where the rows are indexed by directed edges, and define the incidence matrix by

$$M_{(i,j),k} = \begin{cases} +1 & \text{if } k = i \\ -1 & \text{if } k = j \\ 0 & \text{if } k \notin \{i, j\} \end{cases} . \tag{1}$$

Lastly, denote $I_{ij} \in \{0,1\}^{2 \times 2m}$ a matrix whose rows are the columns of the identity matrix $I \in \{0,1\}^{2m \times 2m}$ corresponding to the directed edges (i, j) and (j, i).

$$I_{ij}^T M x(k) = 0. \tag{2}$$

Assumptions **A1** and **A4** imply that the vector of opinions $x(k)$ must minimize $\|x(k) - x(k-1)\|_2$ with $x(k-1)$ fixed, which when combined with the constraint (2) defines $x(k)$ as satisfying

$$x(k) = \text{argmin}_{x \in \mathbb{R}} \{\|x - x(k-1)\|_2^2 : I_{ij}^T M x = 0\}. \tag{3}$$

From equation (3), it's immediate that if nodes i and j communicate at time t, the new opinion values on the network are given by

$$x_l(k) = \begin{cases} \frac{x_i(k-1)+x_j(k-1)}{2} & \text{if } l \in \{i,j\} \\ x_l(k-1) & \text{if } l \notin \{i,j\} \end{cases}. \tag{4}$$

Equation (4) defines how opinions in the network change when a pair of nodes i and j communicate.

2.2. **Key Definitions and Implications of the Update Rule (4).** Denote $X(0) = X$ a random vector of initial opinions on the network and denote

$$X_{\text{ave}} = \frac{X^T 1_n}{n}, \tag{5}$$

where $1_n \in \{1\}^{n \times 1}$ is the vector of all ones, the average value of the initial opinions.

We consider the following three aggregate quantities over the graph G for measuring the total difference in opinion between the individuals in the graph:

Definition 1. *Denote and define the discordance of the nodes $i \in \mathcal{V}(G)$ with values $X_i(t)$ by*

$$d(X, t|G) = \sum_{i,j=1}^{n} A_{ij}(X_i(t) - X_j(t))^2 = \|MX(t)\|_2^2.$$

Definition 2. *Denote and define the total discordance of the nodes $i \in \mathcal{V}(G)$ with values $X_i(t)$ by*

$$d(X, t) = \sum_{i,j=1}^{n} (X_i(t) - X_j(t))^2.$$

Definition 3. *Denote and define the consensus error of the nodes $i \in \mathcal{V}(G)$ with values $X_i(t)$ by*

$$e(X, t) = \|X(t) - X_{ave}\mathbf{1}\|_2^2.$$

We first show the following simple relationship between the total discordance and the consensus error, which shows that they are measuring the same thing.

Proposition 1. *The total discordance in definition 2 is related to the consensus error 3 by*

$$d(X, t) = 2ne(X, t).$$

Proof. We omit the time t in the following calculation for brevity:

$$\begin{aligned} d(X) &= \sum_{i,j=1}^{n} (X_i - X_j)^2 \\ &= \sum_{i,j=1}^{n} (X_i - X_{\text{ave}})^2 + \sum_{i,j=1}^{n} (X_j - X_{\text{ave}})^2 - 2 \sum_{i,j=1}^{n} (X_i - X_{\text{ave}})(X_j - X_{\text{ave}}) && (6) \\ &= 2ne(X) - 2(nX_{\text{ave}} - nX_{\text{ave}})(nX_{\text{ave}} - nX_{\text{ave}}) && (7) \\ &= 2ne(X). && (8) \end{aligned}$$

where in (6) use $X_i - X_j = X_i - X_{\text{ave}} + X_{\text{ave}} - X_j$ and in (7) we pass the sums inside $(X_i - X_{\text{ave}})(X_j - X_{\text{ave}})$. $\square$

The next proposition shows how the total discordance, or equivalently the total consensus error by proposition 1, changes when an update occurs. We phrase the proposition in terms of an update rule that is a bit more general than in other plaes in this document: we consider update schemes such that if nodes p and q update at time t, then for $\alpha_k + \beta_k = 1$,

$$X_p(k+1) = \alpha_k X_p(k) + \beta_k X_q(k) \qquad X_q(k+1) = \beta_k X_p(k) + \alpha_k X_q(k). \tag{9}$$

α_k gives the fraction of information nodes p and q retain at the update, and β_k gives the fraction of information nodes p and q adopt from the node they are updating with. Alternatively, if we let $0 < \alpha_k \leq 1/2$, we write the (**??**):

$$\begin{aligned} X_p(k+1) &= X_p(k) - \alpha_k \operatorname{sgn}(X_p(k) - X_q(k)) \frac{X_p(k) - X_q(k)}{2} \\ X_q(k+1) &= X_q(k) + \alpha_k \operatorname{sgn}(X_p(k) - X_q(k)) \frac{X_p(k) - X_q(k)}{2} \end{aligned} \tag{10}$$

Proposition 2. *Assume that the collection of nodes $S \subset \mathcal{V}(G)$ are updated at timestep k via the update rule*

$$x_i(k+1) = \begin{cases} \beta(\frac{\sum_{j\in S} x_j(k)}{|S|} - x_i(k)) & \textit{if } i \in S \\ x_i(k) & \textit{otherwise.} \end{cases} \tag{11}$$

Then, the change in error at timestep k is given by

$$e(X, k+1) - e(X, k) = -\frac{\beta(2-\beta)}{|S|} (x^S(k))^T L_{|S|} x^S(k), \tag{12}$$

where $x^S(k)$ denotes the components of $x(k)$ corresponding to the nodes S and where $L_{|S|}$ denotes the unnormalized Laplacian of the complete graph of $|S|$ nodes; i.e.

$$(x^S)^T L_{|S|} x^S = \frac{1}{2} \sum_{i,j \in S} (x_i - x_j)^2 \tag{13}$$

Proof. Write the update $x(k+1) = x(k) + \beta(x^S_{\text{ave}}(k) 1_S - x^S(k))$ where

$$x^S_{\text{ave}}(k) = \frac{\sum_{j\in S} x_j(k)}{|S|}$$

$$(1_S)_i = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{if } i \notin S \end{cases}$$

$$x^S(k) = \begin{cases} x_i(k) & \text{if } i \in S \\ 0 & \text{if } i \notin S. \end{cases}$$

Denote by L the Laplacian matrix of the complete graph. Then,

$$x(k+1)^T L x(k+1) - x(k)^T L x(k) = 2\alpha x(k) L (x^S_{\text{ave}}(k) 1_S - x^S(k)) + \alpha^2 (x^S_{\text{ave}} - x^S)^T L (x^S_{\text{ave}} - x^S). \tag{14}$$

Calculating

$$(L x^S_{\text{ave}}(k) 1_S)_i = \begin{cases} -\sum_{j\in S} x_j(k) & \text{if } i \notin S \\ \frac{n-|S|}{|S|} \sum_{j \in S} x_j(k) & \text{if } i \in S. \end{cases} \tag{15}$$

and

$$(L x^S(k))_i = \begin{cases} -\sum_{j\in S} x_j(k) & \text{if } i \notin S \\ (n-1) x_i(k) - \sum_{j\neq i : j \in S} x_j(k) & \text{if } i \in S. \end{cases} \tag{16}$$

we find

$$L(x_{\mathrm{ave}}(k)1_S - x^S(k)) = \begin{cases} 0 & \text{if } i \notin S \\ \frac{n-|S|}{|S|} \sum_{j\in S} x_j(k) - (n-1)x_i(k) + \sum_{j\neq i: j\in S} x_j(k) & \text{if } i \in S. \end{cases} \tag{17}$$

Therefore,

(18)

$$x^T L(x_{\mathrm{ave}}(k)1_S - x^S(k)) = \frac{n-|S|}{|S|}\left(\sum_{j\in S} x_j(k)\right)^2 - (n-1)\sum_{j\in S} x_j(k)^2 + \sum_{i\in S}\sum_{j\neq i: j\in S} x_j(k)$$

$$= \frac{n}{|S|}\left(\sum_{j\in S} x_j(k)\right)^2 - \left(\sum_{j\in S} x_j(k)\right)^2 - n\sum_{j\in S} x_j(k)^2 + \sum_{j\in S} x_j(k)^2 + \sum_{i\in S}\sum_{j\neq i: j\in S} x_j(k) \tag{19}$$

$$= \frac{n}{|S|}\left(\sum_{j\in S} x_j(k)\right)^2 - \left(\sum_{j\in S} x_j(k)\right)^2 - n\sum_{j\in S} x_j(k)^2 + \left(\sum_{j\in S} x_j(k)\right)^2 \tag{20}$$

$$= \frac{n}{|S|}\left(\sum_{j\in S} x_j(k)\right)^2 - n\sum_{j\in S} x_j(k)^2 \tag{21}$$

(22)

Calculating,

$$x^S_{\mathrm{ave}} 1_S L x^S_{\mathrm{ave}} 1_S = \frac{n-|S|}{|S|}\left(\sum_{j\in S} x_j(k)\right)^2 \tag{23}$$

and

$$(x^S)^T L x^S_{\mathrm{ave}} 1_S = \frac{n-|S|}{|S|}\left(\sum_{j\in S} x_j(k)\right)^2 \tag{24}$$

and

$$(x^S)^T L(x^S) = (n-1)\sum_{i\in S} x_i(k)^2 - \sum_{i\in S}\sum_{j\neq i: j\in S} x_i(k)x_j(k) \tag{25}$$

we find

(26)

$$(x^S_{\mathrm{ave}}1_S - x^S)L(x^S_{\mathrm{ave}}1_S - x^S) = -\frac{n}{|S|}\left(\sum_{j\in S} x_j(k)\right)^2 + \left(\sum_{j\in S} x_j(k)\right)^2 + (n-1)\sum_{i\in S} x_i(k)^2 - \sum_{i\in S}\sum_{j\neq i: j\in S} x_i(k)x_j(k)$$

$$= -\frac{n}{|S|}\left(\sum_{j\in S} x_j(k)\right)^2 + n\sum_{i\in S} x_i(k)^2.$$

Therefore,

$$x(k+1)^T L x(k+1) - x(k)^T L x(k) \tag{27}$$

$$= -n\beta(2-\beta)\left(\sum_{j\in S} x_j(k)^2 - \frac{1}{|S|}\left(\sum_{j\in S} x_j(k)\right)^2\right) = -n\frac{\beta(2-\beta)}{|S|}(x^S)^T L_{|S|} x^S.$$

Equation (12) follows from (1). □

The following corollary is the basis for how we study the convergence to consensus of the opinion dynamics process.

Corollary 1. The consensus error satisfies the recurrence relation

$$e(X,k+1) = \left(1 - \frac{\beta(2-\beta)(x^S(k))^T L_{|S|} x^S(k)}{|S| e(X,k)}\right) e(X,k). \tag{28}$$

Denote

$$\alpha(x,\beta,S) = \frac{\beta(2-\beta)(x^S(k))^T L_{|S|} x^S(k)}{|S| e(X,k)} \tag{29}$$

We note that the recurrence is monotonically decreasing:

Proposition 3. *Let M be the first time such that $e_M(X) = 0$, where $M = \infty$ if $e_k(X) > 0$ for all k. The coefficient $\alpha(x,\beta,S)$ from* (1) *satisfies $0 \le \alpha(x,\beta,S) < 1$ for all $k = 0,\dots,M-1$.*

Proof. We have the following inequality

$$e(X,k) = e(X,k+1) + \frac{\beta(2-\beta)}{|S|}(x^S(k))^T L_{|S|} x^S(k) > \frac{\beta(2-\beta)}{|S|}(x^S(k))^T L_{|S|} x^S(k) \tag{30}$$

so that

$$1 > \frac{\beta(2-\beta)(x^S(k))^T L_{|S|} x^S(k)}{|S| e(X,k)} = \alpha(x,\beta,|S|). \tag{31}$$

Certainly $0 \le \alpha(x,\beta,|S|)$. □

(1) An example of when $e_k(X) = 0$ eventually is for a circle graph of an even number of nodes with the greedy update procedure.

Corollary 1 tells us how the consensus error changes at a particular timestep given that a certain group of nodes S updates to consensus. In order to discuss convergence of the opinion dynamics model to consensus, it is required that we specify how a group of nodes is chosen to update.

2.3. **Communication Rates Between Nodes and Convergence to Consensus.** Denote $\mathcal{C}(G)$ the collection of sets of vertices that form connected subgraphs of G. Denote $U(k) \mapsto \mathcal{C}(G)$ the random variable at timestep k determining the collection of vertices to update. Assume that $U(k)$ is independent of $X(l)$ for $l < k-1$ given $X(k)$. Denote $\mathrm{E}_{U(k)}[\cdot]$ the expectation over the random variable $U(k)$. Let $Y \in \mathcal{R}_k$ be a random variable that is a possible value of $X(k)$ at time k.

Denote by $\mathcal{O}(U,\beta)$ the opinion dynamics stochastic process of the update rule (**??**) with the update random variable $U(k)$. We study convergence of the opinion dynamics model $\mathcal{O}(U,\beta)$ by consider the following definition, which defines a notion of an upper bound on the convergence time.

Definition 4. *For all $0 < \epsilon < 1$, the $\epsilon-$averaging time of $\mathcal{O}(U,\beta)$ is denoted by $T_{ave}(\epsilon, \mathcal{O}(U,\beta))$ and defined by*

$$T_{ave}(\epsilon,\mathcal{O}(U,\beta)) = \sup_{x(0)\in\mathbb{R}} \inf_{k\in\mathbb{Z}_+} \left\{ \Pr\left(\frac{\|X(k) - X_{ave}1_n\|}{\|X(0) - X_{ave}1_n\|} \ge \epsilon\right) \le \epsilon \right\} \tag{32}$$

Additionally, we make the following definition, which defines a lower bound on the convergence time of $\mathcal{O}(U,\beta)$.

Definition 5. *For all $0<\epsilon<1$, the best $\epsilon-$averaging time of $\mathcal{O}(U,\beta)$ is denoted by $\mathcal{T}_{ave}(\epsilon,\mathcal{O}(U,\beta))$ and defined by*

$$\mathcal{T}_{ave}(\epsilon,\mathcal{O}(U,\beta)) = \inf_{x(0)\in\mathbb{R}} \inf_{k\in\mathbb{Z}_+} \left\{ \Pr\left(\frac{\|x(k)-x_{ave}1_n\|}{\|x(0)-x_{ave}1_n\|} \geq \epsilon \right) \leq \epsilon \right\} \tag{33}$$

With definitions 4 and 5 in mind, we state the following theorem

Theorem 1. *Assume that $U=U(k)$ and $\beta=\beta(k)$ are independent of k. Denote and define the following quantities*

$$\alpha^*(\beta) = \sup_{Y\in\mathbb{R}^n} \mathrm{E}_{U=S}\left[\alpha(Y,\beta,S)\right] \qquad \alpha_*(\beta) = \inf_{\substack{Y\in\mathbb{R}^n\\ Y\perp 1_n}} \mathrm{E}_{U=S}\left[\alpha(Y,\beta,S)\right] \tag{34}$$

Let $\alpha>0$. The ϵ-averaging time and the best ϵ-averaging time are bounded as follows

$$T_{ave}(\epsilon,\mathcal{O}(U,\beta)) \leq \frac{3\log(\epsilon)}{\log(1-\alpha_*(\beta))} \tag{35}$$

$$\mathcal{T}_{ave}(\epsilon,\mathcal{O}(U,\beta)) \geq \frac{\log(\epsilon)}{\log(1-\alpha^*(\beta))}. \tag{36}$$

Theorem 1 is a result of the following bound on the expectation of the consensus error:

Lemma 1. Assume that $U=U(k)$ and $\beta=\beta(k)$ are independent of k. The expectation $\mathrm{E}_U\left[e(X,k)\right]$, which denotes the expectation over the random variables $U(1),U(2),\ldots,U(k)$ of the consensus error at timestep k, satisfying the following bound

$$(1-\alpha^*(\beta))\,\mathrm{E}_U[e(X,k)] \leq \mathrm{E}_U[e(X,k+1)] \leq (1-\alpha_*(\beta))\,\mathrm{E}_U[e(X,k)]. \tag{37}$$

Proof. From corollary 1, we know that

$$\begin{aligned} \mathrm{E}_{U(k+1)=S}[e(X,k+1)|X(k)=Y] &= e(Y,k) - \mathrm{E}_{U(k+1)=S}\left[\frac{\beta(2-\beta)}{|S|}(Y^S)^T L_{|S|} Y^S\right] \\ &= \left(1-\mathrm{E}_{U(k+1)=S}\left[\frac{\beta(2-\beta)(Y^S)^T L_{|S|} Y^S}{|S| e(Y,k)}\right]\right) e(Y,k), \end{aligned} \tag{38}$$

Thus, using that $\alpha(Y,\beta,S)=\alpha(Y+c1_n,\beta,S)$ for all $c\in\mathbb{R}$, by (38)

$$(1-\alpha^*(\beta))\,e(Y,k) \leq \mathrm{E}_{U(k+1)=S}\left[e(X,k+1)|X(k)=Y\right] \leq (1-\alpha_*(\beta))\,e(Y,k). \tag{39}$$

Taking the expectation over $U(1),\ldots,U(k)$ of both sides of (39) results in (37). □

Proof of theorem 1. By Markov's inequality, lemma 1, and the fact that $e(X,k)=e(X+\beta 1_n,k)$ for all $\beta\in\mathbb{R}$,

$$\Pr\left(\frac{\sqrt{e(X,k)}}{\sqrt{e(X,0)}} \geq \epsilon\right) \leq \frac{1}{\epsilon^2 e(X,0)}\,\mathrm{E}_U[e(X,k)] \leq \epsilon^{-2}(1-\alpha_*(\beta))^k \tag{40}$$

Thus, the upper bound on the ϵ-averaging time in (35) is proven by taking applying $\log(\cdot)$ to both sides of (40).

Since $e(X,k)\leq e(X,0)$,

$$\Pr\left(\frac{\sqrt{e(X,k)}}{\sqrt{e(X,0)}} > \epsilon\right) \geq \frac{\mathrm{E}[e(X,k)]-\epsilon^2 e(X,0)}{e(X,0)-\epsilon^2 e(X,0)} \tag{41}$$

for all $0 < \epsilon < e(X,0)$. Additionally, by lemma 1 and (41)

$$\Pr\left(\frac{\sqrt{e(X,k)}}{\sqrt{e(X,0)}} > \epsilon\right) \geq \frac{(1-\alpha^*(\beta))^k - \epsilon^2}{1-\epsilon^2}. \tag{42}$$

Thus, we see that if

$$k \leq \frac{\log(\epsilon) + \log(1+\epsilon-\epsilon^2)}{\log(1-\alpha^*(\beta))} \leq \frac{\log(\epsilon)}{\log(1-\alpha^*(\beta))}. \tag{43}$$

then

$$\Pr\left(\frac{\sqrt{e(X,k)}}{\sqrt{e(X,0)}} > \epsilon\right) \geq \frac{(1-\alpha^*(\beta))^k - \epsilon^2}{1-\epsilon^2} \geq \epsilon, \tag{44}$$

which is the lower bound we needed to prove. □

3. Insights into Group Updating

In this section, we use the results from previous sections to gain insight into the effects of group updates on reaching convergence.

The following corollary considers the effect of increasing the size of the group being updated on convergence to consensus.

Corollary 2. Suppose that $U(k) = U$ and $\beta(k) = \beta$ are independent of k. Further assume that $|U| = k$ is fixed; i.e., we assume that U may only return subgraphs of a fixed size k. Construct the matrix $P \in [0,1]^{n\times n}$ of probabilities

$$P_{ij} = \Pr((i,j) \in U). \tag{45}$$

and let $Q \in [0,1]^{2m\times 2m}$ be a diagonal matrix with $Q_{(i,j)} = \sqrt{P_{ij}}$. Additionally, denote $M \in \{-1,0,1\}^{2m\times n}$ the incidence matrix of the graph G and let W be the matrix

$$W(\beta) = I - \frac{\beta(2-\beta)}{2k}(QM)^*QM. \tag{46}$$

Denote $\sigma_k(\cdot)$ the k^{th} singular value in decreasing order of matrix and denote $\lambda_k(\cdot)$ the k^{th} eigenvalue in decreasing order of a matrix. We find that

$$1-\alpha_*(\beta) = 1 - \frac{\beta(2-\beta)}{2k}\sigma_{n-1}(QM)^2 = \lambda_2(W(\beta)) \qquad 1-\alpha^*(\beta) = 1 - \frac{\beta(2-\beta)}{2k}\sigma_1(QM)^2 = \lambda_n(W(\beta)) \tag{47}$$

and

$$T_{\text{ave}}(\epsilon, \mathcal{O}(U,\beta)) \leq \frac{3\log(\epsilon)}{\log(\lambda_2(W(\beta)))} = \frac{3\log(\epsilon)}{\log(1-\frac{\beta(2-\beta)}{2k}\sigma_{n-1}(QM)^2)} \tag{48}$$

$$\mathcal{T}_{\text{ave}}(\epsilon, \mathcal{O}(U,\beta)) \geq \frac{\log(\epsilon)}{\log(\lambda_n(W(\beta)))} = \frac{\log(\epsilon)}{\log(1-\frac{\beta(2-\beta)}{2k}\sigma_1(QM)^2)}. \tag{49}$$

Additionally, we have the following recursive bound on $\mathrm{E}_U[e(X,k)]$:

$$\begin{aligned}(1-\frac{1}{2}\sigma_1(QM)^2)\,\mathrm{E}_U[e(X,k)] &= \lambda_n(W)\,\mathrm{E}_U[e(X,k)] \\ &\leq \mathrm{E}_U[e(X,k+1)] \leq \lambda_2(W)\,\mathrm{E}_U[e(X,k)] = (1-\frac{1}{2}\sigma_{n-1}(QM)^2)\,\mathrm{E}_U[e(X,k)]\end{aligned} \tag{50}$$

Proof. The corollary is a result of the following observations. The first is that

$$\frac{\mathrm{E}_{U=S}\left[(Y^S)^T L_{|S|} Y^S\right]}{e(Y,k)} = \frac{\|QMY\|_2^2}{2\|Y\|_2^2} \tag{51}$$

is a Rayleigh quotient. The second is that

$$1 - \frac{\|QMY\|_2^2}{\|Y\|_2^2} = \frac{Y^T(I-(QM)^*QM)Y}{\|Y\|_2^2} = \frac{Y^TWY}{\|Y\|_2^2}. \tag{52}$$

$\square$

References