Supplementary Material of "Formalizing Word Sampling for Vocabulary Prediction as Graph-based Active Learning"

This supplementary material supplements the proof of Theorem 3.1. We let $\mathbf{M} = \mathbf{\Lambda}$.

1 Lemma 1

In Lemma 1, we prove that $\mathbf{u}_i^\top \mathbf{H}_{k+1}^{-1} \mathbf{u}_i$, $\mathbf{u}_i^\top \mathbf{H}_{k+1}^{-1} \mathbf{u}_j$, and $\left(\mathbf{H}_{k+1}^{-1} \mathbf{u}_i\right)^\top \mathbf{M}^{-1} \left(\mathbf{H}_{k+1}^{-1} \mathbf{u}_i\right)$ is constant over the choice of i and depend only on k. We prove this by induction. We let $j \neq i$. When $k = 0$, for every i, j,

$$\left(\mathbf{H}_0^{-1}\mathbf{u}_i\right)^\top \Lambda^{-1} \left(\mathbf{H}_0^{-1}\mathbf{u}_i\right) = \tfrac{n-1}{n} \left(\mu^2 \left(\tfrac{n}{n-1}\right)^2 + \mu \left(\tfrac{n}{n-1}\right)\right)^4 \mathbf{q}_i^\top \mathbf{q}_i + \epsilon \tfrac{1}{n} \tag{1}$$

$$= \quad \tfrac{n-1}{n} \left(\mu^2 \left(\tfrac{n}{n-1}\right)^2 + \mu \left(\tfrac{n}{n-1}\right)\right)^4 \tfrac{n-1}{n} + \epsilon \tfrac{1}{n} = C_0^{\text{numera}} \tag{2}$$

$$\left(\mathbf{H}_0^{-1}\mathbf{u}_i\right)^\top \Lambda^{-1} \left(\mathbf{H}_0^{-1}\mathbf{u}_i\right) = \tfrac{n-1}{n} \left(\mu^2 \left(\tfrac{n}{n-1}\right)^2 + \mu \left(\tfrac{n}{n-1}\right)\right)^4 \mathbf{q}_i^\top \mathbf{q}_j + \epsilon \tfrac{1}{n} \tag{3}$$

$$= \quad \tfrac{n-1}{n} \left(\mu^2 \left(\tfrac{n}{n-1}\right)^2 + \mu \left(\tfrac{n}{n-1}\right)\right)^4 \left(-\tfrac{1}{n}\right) + \epsilon \tfrac{1}{n} = C_0^{\text{nutero}} \tag{4}$$

$$\mathbf{u}_i^\top \mathbf{H}_0^{-1} \mathbf{u}_i = \left(\mu^2 \left(\tfrac{n}{n-1}\right)^2 + \mu \left(\tfrac{n}{n-1}\right)\right)^2 \mathbf{q}_i^\top \mathbf{q}_i + \epsilon \tfrac{1}{n} = \left(\mu^2 \left(\tfrac{n}{n-1}\right)^2 + \mu \left(\tfrac{n}{n-1}\right)\right)^2 \tfrac{n-1}{n} + \epsilon \tfrac{1}{n} = C_0^{\text{pure}} \tag{5}$$

$$\mathbf{u}_i^\top \mathbf{H}_0^{-1} \mathbf{u}_j = \left(\mu^2 \left(\tfrac{n}{n-1}\right)^2 + \mu \left(\tfrac{n}{n-1}\right)\right)^2 \mathbf{q}_i^\top \mathbf{q}_j + \epsilon \tfrac{1}{n} = \left(\mu^2 \left(\tfrac{n}{n-1}\right)^2 + \mu \left(\tfrac{n}{n-1}\right)\right)^2 \left(-\tfrac{1}{n}\right) + \epsilon \tfrac{1}{n} = C_0^{\text{cross}} \tag{6}$$

Thus, these terms are constant over i, j.

Assume that the four terms are constant at k. Then, when $k+1$,

$$\begin{aligned}
& \mathbf{u}_i^\top \mathbf{H}_{k+1}^{-1} \mathbf{u}_j \\
= \quad & \mathbf{u}_i^\top \left(\mathbf{H}_k^{-1} - \frac{\left(\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right)\left(\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right)^\top}{1 + \mathbf{u}_{i_{k+1}}^\top \mathbf{H}_k^{-1} \mathbf{u}_{i_{k+1}}} \right) \mathbf{u}_j \\
= \quad & \mathbf{u}_i^\top \mathbf{H}_k^{-1} \mathbf{u}_j - \mathbf{u}_i^\top \frac{\left(\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right)\left(\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right)^\top}{1 + \mathbf{u}_{i_{k+1}}^\top \mathbf{H}_k^{-1} \mathbf{u}_{i_{k+1}}} \mathbf{u}_j \\
= \quad & C_k^{\text{cross}} - \frac{\left(\mathbf{u}_i^\top \mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right)\left(\mathbf{u}_{i_{k+1}}^\top \mathbf{H}_k^{-1}\mathbf{u}_j\right)}{1 + \mathbf{u}_{i_{k+1}}^\top \mathbf{H}_k^{-1} \mathbf{u}_{i_{k+1}}} \\
= \quad & C_k^{\text{cross}} - \frac{\left(C_k^{\text{cross}}\right)^2}{1 + C_k^{\text{pure}}}
\end{aligned} \tag{7}$$

$$\mathbf{u}_i^\top \mathbf{H}_{k+1}^{-1}\mathbf{u}_i = C_k^{\text{pure}} - \frac{\left(C_k^{\text{cross}}\right)^2}{1+C_k^{\text{pure}}} \tag{8}$$

$$\begin{aligned}
& C_{k+1}^{\text{nutero}} = \left(\mathbf{H}_{k+1}^{-1}\mathbf{u}_i\right)^\top \mathbf{M}^{-1}\left(\mathbf{H}_{k+1}^{-1}\mathbf{u}_j\right) && (9)\\
= \; & \left(\mathbf{H}_k^{-1}\mathbf{u}_i - \frac{C_k^{\text{cross}}}{1+C_k^{\text{pure}}}\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right)^\top \mathbf{M}^{-1}\left(\mathbf{H}_k^{-1}\mathbf{u}_j - \frac{C_k^{\text{cross}}}{1+C_k^{\text{pure}}}\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right) && (10)\\
= \; & \left(\mathbf{H}_k^{-1}\mathbf{u}_i\right)^\top \mathbf{M}^{-1}\left(\mathbf{H}_k^{-1}\mathbf{u}_j\right) - \left(\mathbf{H}_k^{-1}\mathbf{u}_i\right)\mathbf{M}^{-1}\left(\frac{C_k^{\text{cross}}}{1+C_k^{\text{pure}}}\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right) && (11)\\
- \; & \frac{C_k^{\text{pure}}}{1+C_k^{\text{pure}}}\left(\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right)^\top \mathbf{M}^{-1}\left(\mathbf{H}_k^{-1}\mathbf{u}_i\right) + \left(\frac{C_k^{\text{cross}}}{1+C_k^{\text{pure}}}\right)^2\left(\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right)^\top \mathbf{M}^{-1}\left(\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right) && (12)\\
= \; & C_k^{\text{nutero}} - \frac{2C_k^{\text{cross}}}{1+C_k^{\text{pure}}}C_k^{\text{nutero}} + \left(\frac{C_k^{\text{cross}}}{1+C_k^{\text{pure}}}\right)^2 C_k^{\text{numera}} && (13)
\end{aligned}$$

$$\begin{aligned}
& C_{k+1}^{\text{numera}} = \left(\mathbf{H}_{k+1}^{-1}\mathbf{u}_i\right)^\top \mathbf{M}^{-1}\left(\mathbf{H}_{k+1}^{-1}\mathbf{u}_i\right) && (14)\\
= \; & \left(\mathbf{H}_k^{-1}\mathbf{u}_i - \tfrac{C_k^{\text{cross}}}{1+C_k^{\text{pure}}}\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right)^\top \mathbf{M}^{-1}\left(\mathbf{H}_k^{-1}\mathbf{u}_i - \tfrac{C_k^{\text{cross}}}{1+C_k^{\text{pure}}}\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right) && (15)\\
= \; & \left(\left(\mathbf{H}_k^{-1}\mathbf{u}_i\right)^\top - \tfrac{C_k^{\text{cross}}}{1+C_k^{\text{pure}}}\left(\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right)^\top\right)\mathbf{M}^{-1}\left(\left(\mathbf{H}_k^{-1}\mathbf{u}_i\right) - \tfrac{C_k^{\text{cross}}}{1+C_k^{\text{pure}}}\left(\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right)\right) && (16)\\
= \; & \left(\mathbf{H}_k^{-1}\mathbf{u}_i\right)^\top \mathbf{M}^{-1}\left(\mathbf{H}_k^{-1}\mathbf{u}_i\right) - \left(\mathbf{H}_k^{-1}\mathbf{u}_i\right)\mathbf{M}^{-1}\left(\tfrac{C_k^{\text{cross}}}{1+C_k^{\text{pure}}}\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right) && (17)\\
- \; & \tfrac{C_k^{\text{cross}}}{1+C_k^{\text{pure}}}\left(\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right)^\top \mathbf{M}^{-1}\left(\mathbf{H}_k^{-1}\mathbf{u}_i\right) + \left(\tfrac{C_k^{\text{cross}}}{1+C_k^{\text{pure}}}\right)^2\left(\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right)^\top \mathbf{M}^{-1}\left(\mathbf{H}_k^{-1}\mathbf{u}_{i_{k+1}}\right) && (18)\\
= \; & C_k^{\text{numera}} - \tfrac{2C_k^{\text{cross}}}{1+C_k^{\text{pure}}}C_k^{\text{nutero}} + \left(\tfrac{C_k^{\text{cross}}}{1+C_k^{\text{pure}}}\right)^2 C_k^{\text{numera}} && (19)
\end{aligned}$$

Thus, the four terms are constant at $k+1$ if they are constant at k. Because they are constant when $k = 0$, we proved they are constant for every $k \in \{0, 1, 2, \ldots,\}$.

2 Lemma 2

Next, we will prove Lemma 2. Lemma 2 states that $\text{score}(k, i) - \text{score}(k+1, i) > 0$ for every $k \in \{0, 1, 2, \ldots, n-1\}$ if the graph concerned is a complete graph. Because Lemma 1 proves that, for a complete graph, $\text{score}(k, i)$ is constant over i at each round k, we only need to prove that the constant for the score at the k-th round is always greater than that for the score at the $k+1$-th round.

Our proof strategy is simple: we repeat factorizing $\text{score}(k, i) - \text{score}(k+1, i)$ into a factor that can be proved to be positive and the remaining factor. Then, because we only care about the sign of $\text{score}(k, i) - \text{score}(k+1, i)$, we only need to focus on the sign of the remaining factor. Finally, we can eliminate k from the remaining factor and we can also prove the remaining factor to be positive under some conditions on μ, n, and ϵ that can exclude exceptional cases.

Remember that the score can be written as follows using the constants that we defined to prove Lemma 1:

$$\text{score}(k, i) = \frac{C_k^{\text{numera}}}{1+C_k^{\text{pure}}}. \tag{20}$$

2.1 Sign of Constants

First, we prove the sign of the four types of the constants that we introduced to prove Lemma 1. Namely, we can prove that $C_k^{\text{pure}} > 0$, $C_k^{\text{numera}} > 0$, $C_k^{\text{cross}} < 0$, $C_k^{\text{nutero}} < 0$ for every k.

2.1.1 Proof of $C_k^{\textbf{pure}} > 0$

By definition, $C_k^{\text{pure}} \stackrel{\text{def}}{=} \mathbf{u}_i^\top \mathbf{H}_k^{-1} \mathbf{u}_i$. Thus, if $\mathbf{H}_k^{-1}$ is *positive definite*, i.e., for all non-zero vector $\mathbf{x}$, $\mathbf{x}^\top \mathbf{H}_k^{-1} \mathbf{x} > 0$ holds, then, we can prove $C_k^{\text{pure}} = \mathbf{u}_i^\top \mathbf{H}_k^{-1} \mathbf{u}_i > 0$. Obviously, $\mathbf{u}_i$ is a non-zero vector for every i since $\mathbf{u}_i$ is, by definition, an orthonormal vector, which $||\mathbf{u}_i||_2 = 1$.

We show that $\mathbf{H}_k^{-1}$ is positive definite as follows. From [1], by definition,

$$\mathbf{H}_k \stackrel{\text{def}}{=} \Gamma^{-1} + \mathbf{U}_{\mathcal{L}_k}^\top \mathbf{U}_{\mathcal{L}_k}. \tag{21}$$

Then, by multiplying (21) by $\mathbf{x}^\top$ and $\mathbf{x}$ from the left and right side, respectively, we obtain the following form.

$$\mathbf{x}^\top \mathbf{H}_k \mathbf{x} = \mathbf{x}^\top \Gamma^{-1} \mathbf{x} + \left(\mathbf{U}_{\mathcal{L}_k} \mathbf{x}\right)^\top \left(\mathbf{U}_{\mathcal{L}_k} \mathbf{x}\right) \tag{22}$$

In (22), Γ^{-1} in the first term in the right hand side is positive definite because Γ^{-1} is, by definition, a diagonal matrix whose diagonal elements are all positive, and thus, all of its eigenvalues are positive as follows. Note that, by definition, $\mu > 0$, $\epsilon > 0$, and $n \geq 2 > 0$.

$$\Gamma = \operatorname{diag}\left(\left(\left(\mu^2 \left(\frac{n}{n-1}\right)^2 + 2\mu \left(\frac{n}{n-1}\right)\right)^{-1}, \ldots, \left(\mu^2 \epsilon^2 + 2\mu\epsilon\right)^{-1}\right)^\top\right). \tag{23}$$

Next, in (22), the second term in the right hand side, obviously $\left(\mathbf{U}_{\mathcal{L}_k} \mathbf{x}\right)^\top \left(\mathbf{U}_{\mathcal{L}_k} \mathbf{x}\right) \geq 0$. Thus, $\mathbf{U}_{\mathcal{L}_k}^\top \mathbf{U}_{\mathcal{L}_k}$ is *positive semi-definite*.

Thus, (22) is the sum of a positive definite matrix and a positive semi-definite matrix, which is positive definite. Thus, $\mathbf{H}_k$ is positive definite. Finally, since the inverse of a positive definite matrix is also positive definite, $\mathbf{H}_k^{-1}$ is positive definite.

$\square$

2.1.2 Proof of $C_k^{\textbf{numera}} > 0$

By definition,

$$C_k^{\text{numera}} \stackrel{\text{def}}{=} \mathbf{u}_i^\top \mathbf{H}_k^{-1} \mathbf{\Lambda}^{-1} \mathbf{H}_k^{-1} \mathbf{u}_i = \mathbf{u}_i^\top \left(\mathbf{H}_k^{-1} \mathbf{\Lambda}^{-1} \mathbf{H}_k^{-1}\right) \mathbf{u}_i. \tag{24}$$

Thus, if $\left(\mathbf{H}_k^{-1} \mathbf{\Lambda}^{-1} \mathbf{H}_k^{-1}\right)$ is positive definite, $C_k^{\text{numera}} > 0$. Λ^{-1}, a diagonal matrix, is positive definite because its all diagonal elements are positive by definition as follows:

$$\Lambda^{-1} = \operatorname{diag}\left(\left(\frac{n}{n-1}\right)^{-1}, \ldots, \epsilon^{-1}\right). \tag{25}$$

We proved that $\mathbf{H}_k^{-1}$ is positive definite in §2.1.1. For two positive matrices of the same size, A and B, ABA and BAB are positive definite. Thus, since $\mathbf{H}_k^{-1}$ and Λ^{-1} are positive definite, $\mathbf{H}_k^{-1} \mathbf{\Lambda}^{-1} \mathbf{H}_k^{-1}$ is positive definite.

$\square$

2.1.3 Proof of $C_k^{\text{cross}} < 0$

We prove by induction. When $k = 0$, $C_0^{\text{cross}} < 0$ from (6). Assume $C_k^{\text{cross}} < 0$ at the k-th round. Then, C_{k+1}^{cross} can be written as:

$$C_{k+1}^{\text{cross}} = C_k^{\text{cross}} - \frac{\left(C_k^{\text{cross}}\right)^2}{1 + C_k^{\text{pure}}}. \tag{26}$$

In the right hand side of (26), the first term, $C_k^{\text{cross}} < 0$ by assumption, and the second term $-\frac{\left(C_k^{\text{cross}}\right)^2}{1+C_k^{\text{pure}}} < 0$ since we proved that $C_k^{\text{pure}} > 0$ for every k. Thus, C_{k+1}^{cross}. Thus, by induction, for every k, $C_k^{\text{cross}} < 0$.

□

2.2 Equivalent Form

Now, we go back to proving $\text{score}(k, i) - \text{score}(k + 1, i) > 0$. Since we proved that $C_k^{\text{pure}} > 0$ for every k,

$$\text{score}\,(k, i) - \text{score}\,(k + 1, i) = \frac{C_k^{\text{numera}}}{1 + C_k^{\text{pure}}} - \frac{C_{k+1}^{\text{numera}}}{1 + C_{k+1}^{\text{pure}}} > 0 \tag{27}$$

$$\Leftrightarrow \quad \left(1 + C_{k+1}^{\text{pure}}\right) C_k^{\text{numera}} - C_{k+1}^{\text{numera}} \left(1 + C_k^{\text{pure}}\right) > 0. \tag{28}$$

Thus, (28) is equivalent to Lemma 2.

(28) involves both k and $k + 1$. Since this is intricate to handle, we transform (28) to an equivalent proposition that involves only k as follows.

$$\begin{aligned}
& \left(1 + C_{k+1}^{\text{pure}}\right) C_k^{\text{numera}} - \left(1 + C_k^{\text{pure}}\right) C_{k+1}^{\text{numera}} & (29)\\
= & \left(1 + \left(C_k^{\text{pure}} - \frac{\left(C_k^{\text{cross}}\right)^2}{1 + C_k^{\text{pure}}}\right)\right) C_k^{\text{numera}} - \left(1 + C_k^{\text{pure}}\right) \left(C_k^{\text{numera}} - \frac{2C_k^{\text{cross}}}{1 + C_k^{\text{pure}}} C_k^{\text{nutero}} + \left(\frac{C_k^{\text{cross}}}{1 + C_k^{\text{pure}}}\right)^2 C_k^{\text{numera}}\right) & \\
= & \left(1 + C_k^{\text{pure}}\right) C_k^{\text{numera}} - \left(1 + C_k^{\text{pure}}\right) C_k^{\text{numera}} + 2C_k^{\text{cross}} C_k^{\text{nutero}} - \frac{2\left(C_k^{\text{cross}}\right)^2}{1 + C_k^{\text{pure}}} C_k^{\text{numera}} & (30)\\
= & \ 2C_k^{\text{cross}} C_k^{\text{nutero}} - \frac{2\left(C_k^{\text{cross}}\right)^2}{1 + C_k^{\text{pure}}} C_k^{\text{numera}} & (31)\\
= & \left(-\frac{2C_k^{\text{cross}}}{1 + C_k^{\text{pure}}}\right) \left(C_k^{\text{cross}} C_k^{\text{numera}} - C_k^{\text{nutero}} \left(1 + C_k^{\text{pure}}\right)\right). & (32)
\end{aligned}$$

In (32), since $C_k^{\text{pure}} > 0$ and $C_k^{\text{cross}} < 0$ for every k, its first factor $-\frac{2C_k^{\text{cross}}}{1+C_k^{\text{pure}}} > 0$. Thus, the second factor of (32) must be positive if and only if $C_k^{\text{cross}} C_k^{\text{numera}} - C_k^{\text{nutero}} \left(1 + C_k^{\text{pure}}\right) > 0$:

$$\left(1 + C_{k+1}^{\text{pure}}\right) C_k^{\text{numera}} - \left(1 + C_k^{\text{pure}}\right) C_{k+1}^{\text{numera}} \quad > \quad 0 \tag{33}$$

$$\Leftrightarrow C_k^{\text{cross}} C_k^{\text{numera}} - C_k^{\text{nutero}} \left(1 + C_k^{\text{pure}}\right) \quad > \quad 0 \tag{34}$$

Thus, Lemma 2 is equivalent to (34).

2.3 Removing k-dependent Constants

Since we showed that Lemma 2 is equivalent to (34), we prove (34), which we restate in (35).

$$C_k^{\text{cross}} C_k^{\text{numera}} - C_k^{\text{nutero}} \left(1 + C_k^{\text{pure}}\right) > 0 \tag{35}$$

(35) involves four k-dependent constants: namely, C_k^{cross}, C_k^{numera}, C_k^{pure}, and C_k^{nutero}. These constants are indexed by k and thus depend on k. In this subsection, rewrite (35) by replacing these k-dependent constants with k-free constants that do not depend on k.

Namely, we introduce the following k-free constants as follows:

$$\begin{aligned}
C^{\text{Gpure}} &\stackrel{\text{def}}{=} \mathbf{u}_i^\top \Gamma \mathbf{u}_i = \left(\mu^2 \left(\frac{n}{n-1}\right)^2 + 2\mu \left(\frac{n}{n-1}\right)\right) \left(\frac{n-1}{n}\right) + \left(\mu^2\epsilon^2 + 2\mu\epsilon\right) \frac{1}{n} && (36)\\
C^{\text{Gcross}} &\stackrel{\text{def}}{=} \mathbf{u}_j^\top \Gamma \mathbf{u}_i = \left(\mu^2 \left(\frac{n}{n-1}\right)^2 + 2\mu \left(\frac{n}{n-1}\right)\right) \left(-\frac{1}{n}\right) + \left(\mu^2\epsilon^2 + 2\mu\epsilon\right) \frac{1}{n} && (37)\\
C^{\text{Gnutero}} &\stackrel{\text{def}}{=} \mathbf{u}_i^\top \Gamma \mathbf{\Lambda}^{-1} \Gamma \mathbf{u}_j \\
&= \left(\mu^2 \left(\frac{n}{n-1}\right)^2 + 2\mu \left(\frac{n}{n-1}\right)\right)^2 \frac{n-1}{n} \left(-\frac{1}{n}\right) + \epsilon^{-1} \left(\mu^2\epsilon^2 + 2\mu\epsilon\right)^2 \frac{1}{n} && (38)\\
C^{\text{Gnumera}} &\stackrel{\text{def}}{=} \mathbf{u}_i^\top \Gamma \mathbf{\Lambda}^{-1} \Gamma \mathbf{u}_i \\
&= \left(\mu^2 \left(\frac{n}{n-1}\right)^2 + 2\mu \left(\frac{n}{n-1}\right)\right)^2 \frac{n-1}{n} \left(\frac{n-1}{n}\right) + \epsilon^{-1} \left(\mu^2\epsilon^2 + 2\mu\epsilon\right)^2 \frac{1}{n} && (39)
\end{aligned}$$

In the above calculation, we used

$$\Gamma \mathbf{u}_i = \left(\left(\mu^2 \left(\frac{n}{n-1}\right)^2 + 2\mu \left(\frac{n}{n-1}\right)\right) \mathbf{q}_i^\top, \left(\mu^2\epsilon^2 + 2\mu\epsilon\right) \frac{-1}{\sqrt{n}}\right)^\top. \tag{40}$$

We decompose $\mathbf{H}_k^{-1}$ by using the Woodbury-identity formula as follows.

$$\begin{aligned}
\mathbf{H}_k^{-1} &= \left(\Gamma^{-1} + \mathbf{U}_{\mathcal{L}_k}^\top \mathbf{U}_{\mathcal{L}_k}\right)^{-1} && (41)\\
&= \Gamma - \Gamma \mathbf{U}_{\mathcal{L}_k}^\top \left(\mathbf{I} + \mathbf{U}_{\mathcal{L}_k} \Gamma \mathbf{U}_{\mathcal{L}_k}^\top\right)^{-1} \mathbf{U}_{\mathcal{L}_k} \Gamma. && (42)
\end{aligned}$$

By using (42), we want to rewrite k-dependent constants with k-free constants. To this end, we introduce two formulae that can be used to the rewrite.

First, by using k-free constants, for every $j \notin \mathcal{L}_k$,

$$\mathbf{U}_{\mathcal{L}_k} \Gamma \mathbf{u}_j = \begin{pmatrix} \mathbf{u}_{l_1}^\top \\ \vdots \\ \mathbf{u}_{l_k}^\top \end{pmatrix} \Gamma \mathbf{u}_j = \begin{pmatrix} \mathbf{u}_{l_1}^\top \Gamma \mathbf{u}_j \\ \vdots \\ \mathbf{u}_{l_k}^\top \Gamma \mathbf{u}_j \end{pmatrix} = \begin{pmatrix} C^{\text{Gcross}} \\ \vdots \\ C^{\text{Gcross}} \end{pmatrix} = \mathbf{1}_k C^{\text{Gcross}}. \tag{43}$$

In (43), $\mathbf{1}_k$ is a *ones vector* of size k, i.e., a k-sized vector whose elements are all 1. Also, in (43), we let $\mathcal{L}_k \stackrel{\text{def}}{=} \{l_1, \ldots, l_k\}$.

Second, we want to calculate $\left(\mathbf{I}+\mathbf{U}_{\mathcal{L}_k}\Gamma\mathbf{U}_{\mathcal{L}_k}^\top\right)^{-1}\mathbf{1}_k$. This can be easily computed since $\mathbf{1}_k$ is an eigenvector of $\left(\mathbf{I}+\mathbf{U}_{\mathcal{L}_k}\Gamma\mathbf{U}_{\mathcal{L}_k}^\top\right)^{-1}$ for every k. To see this, we show that $\mathbf{1}_k$ is an eigenvector of $\left(\mathbf{I}+\mathbf{U}_{\mathcal{L}_k}\Gamma\mathbf{U}_{\mathcal{L}_k}^\top\right)$ as follows:

$$\begin{aligned}
\left(\mathbf{I}+\mathbf{U}_{\mathcal{L}_k}\Gamma\mathbf{U}_{\mathcal{L}_k}^\top\right)\mathbf{1}_k &= \mathbf{I}\mathbf{1}_k + \begin{pmatrix} C^{\text{Gpure}} & C^{\text{Gcross}} & \cdots & C^{\text{Gcross}} \\ C^{\text{Gcross}} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & C^{\text{Gcross}} \\ C^{\text{Gcross}} & \cdots & C^{\text{Gcross}} & C^{\text{Gpure}} \end{pmatrix}\mathbf{1}_k && (44)\\
&= \mathbf{1}_k + \mathbf{1}_k\left(C^{\text{Gpure}} + (k-1)\,C^{\text{Gcross}}\right) && (45)\\
&= \left(1 + C^{\text{Gpure}} + (k-1)\,C^{\text{Gcross}}\right)\mathbf{1}_k. && (46)
\end{aligned}$$

Thus, $\mathbf{1}_k$ is an eigenvector of $\left(\mathbf{I}+\mathbf{U}_{\mathcal{L}_k}\Gamma\mathbf{U}_{\mathcal{L}_k}^\top\right)$, with the eigenvalue $\left(1 + C^{\text{Gpure}} + (k-1)\,C^{\text{Gcross}}\right)$. The eigenvectors of the inverse of a matrix is identical to those of the original matrix except that their corresponding eigenvalues are inverted. Thus, the following holds.

$$\left(\mathbf{I}+\mathbf{U}_{\mathcal{L}_k}\Gamma\mathbf{U}_{\mathcal{L}_k}^\top\right)^{-1}\mathbf{1}_k = \frac{\mathbf{1}_k}{1 + C^{\text{Gpure}} + (k-1)\,C^{\text{Gcross}}} \qquad = \frac{\mathbf{1}_k}{C^{\text{Idenom}}}. \tag{47}$$

Here, we defined C^{Idenom} as follows:

$$C^{\text{Idenom}} \stackrel{\text{def}}{=} 1 + C^{\text{Gpure}} + (k-1)\,C^{\text{Gcross}}. \tag{48}$$

By using (43) and (47), we can rewrite k-dependent constants with k-free constants as follows:

$$\begin{aligned}
C_k^{\text{cross}} &= \mathbf{u}_i^\top\mathbf{H}_k^{-1}\mathbf{u}_j && (49)\\
&= \mathbf{u}_i^\top\Gamma\mathbf{u}_j - \mathbf{u}_i^\top\Gamma\mathbf{U}_{\mathcal{L}_k}^\top\left(\mathbf{I}+\mathbf{U}_{\mathcal{L}_k}\Gamma\mathbf{U}_{\mathcal{L}_k}^\top\right)^{-1}\mathbf{U}_{\mathcal{L}_k}\Gamma\mathbf{u}_j && (50)\\
&= C^{\text{Gcross}} - \left(C^{\text{Gcross}}\right)^2\mathbf{1}_k^\top\left(\mathbf{I}+\mathbf{U}_{\mathcal{L}_k}\Gamma\mathbf{U}_{\mathcal{L}_k}^\top\right)^{-1}\mathbf{1}_k && (51)\\
&= C^{\text{Gcross}} - \left(C^{\text{Gcross}}\right)^2\frac{k}{1 + C^{\text{Gpure}} + (k-1)\,C^{\text{Gcross}}} && (52)\\
&= C^{\text{Gcross}} - \frac{\left(C^{\text{Gcross}}\right)^2 k}{C^{\text{Idenom}}} && (53)
\end{aligned}$$

Note that $\mathbf{u}_i$ and $\mathbf{u}_j$ are NOT in $\mathcal{L}_k$ since they are candidate of the $k+1$-th node. Similarly to (53), we can decompose C_k^{pure} as follows:

$$C^{\text{Gpure}} - \frac{\left(C^{\text{Gcross}}\right)^2 k}{C^{\text{Idenom}}} \tag{54}$$

Then,

$$
\begin{aligned}
C_k^{\text{nutero}} &= \mathbf{u}_i^\top \mathbf{H}_k^{-1} \mathbf{\Lambda}^{-1} \mathbf{H}_k^{-1} \mathbf{u}_j && (55)\\
&= \left(\Gamma \mathbf{u}_i - \Gamma \mathbf{U}_{\mathcal{L}_k}^\top \mathbf{1}_k \left(1 + C^{\text{Gpure}} + (k-1)\, C^{\text{Gcross}}\right)^{-1} C^{\text{Gcross}}\right)^\top \\
&\quad \mathbf{\Lambda}^{-1} \left(\Gamma \mathbf{u}_j - \Gamma \mathbf{U}_{\mathcal{L}_k}^\top \mathbf{1}_k \left(1 + C^{\text{Gpure}} + (k-1)\, C^{\text{Gcross}}\right)^{-1} C^{\text{Gcross}}\right) && (56)\\
&= \mathbf{u}_i^\top \Gamma \mathbf{\Lambda}^{-1} \Gamma \mathbf{u}_j \\
&\quad - \mathbf{1}_k^\top \mathbf{U}_{\mathcal{L}_k} \Gamma \mathbf{\Lambda}^{-1} \Gamma \mathbf{u}_j \left(1 + C^{\text{Gpure}} + (k-1)\, C^{\text{Gcross}}\right)^{-1} C^{\text{Gcross}} \\
&\quad - \mathbf{u}_i^\top \Gamma \mathbf{\Lambda}^{-1} \Gamma \mathbf{U}_{\mathcal{L}_k}^\top \mathbf{1}_k \left(1 + C^{\text{Gpure}} + (k-1)\, C^{\text{Gcross}}\right)^{-1} C^{\text{Gcross}} \\
&\quad + \mathbf{1}_k^\top \mathbf{U}_{\mathcal{L}_k} \Gamma \mathbf{\Lambda}^{-1} \Gamma \mathbf{U}_{\mathcal{L}_k}^\top \mathbf{1}_k \left(\left(1 + C^{\text{Gpure}} + (k-1)\, C^{\text{Gcross}}\right)^{-1} C^{\text{Gcross}}\right)^2 && (57)\\
&= C^{\text{Gnutero}} - \frac{2k C^{\text{Gnutero}} C^{\text{Gcross}}}{1 + C^{\text{Gpure}} + (k-1)\, C^{\text{Gcross}}} + \frac{\left(k C^{\text{Gnumera}} + k\,(k-1)\, C^{\text{Gnutero}}\right)\left(C^{\text{Gcross}}\right)^2}{\left(1 + C^{\text{Gpure}} + (k-1)\, C^{\text{Gcross}}\right)^2} && (58)\\
&= C^{\text{Gnutero}} - \frac{2k C^{\text{Gnutero}} C^{\text{Gcross}}}{C^{\text{Idenom}}} + \frac{\left(k C^{\text{Gnumera}} + k\,(k-1)\, C^{\text{Gnutero}}\right)\left(C^{\text{Gcross}}\right)^2}{\left(C^{\text{Idenom}}\right)^2}. && (59)
\end{aligned}
$$

Likewise, C_k^{numera} can be decomposed as follows:

$$
C_k^{\text{numera}} = C^{\text{Gnumera}} - \frac{2k C^{\text{Gnutero}} C^{\text{Gcross}}}{\left(C^{\text{Idenom}}\right)} + \frac{\left(k C^{\text{Gnumera}} + k\,(k-1)\, C^{\text{Gnutero}}\right)\left(C^{\text{Gcross}}\right)^2}{\left(C^{\text{Idenom}}\right)^2}. \qquad (60)
$$

2.4 k-free Equivalent Form

We can rewrite (35) with k-free constants by substituting (53), (54), (59), and (60) to it. However, a simple substitution makes the formula too intricate to write down. Instead, in this subsection, we obtain an equivalent proposition of Lemma 2 that is totally free from k.

To this end, we show that $C^{\text{Idenom}} > 0$.

$$
\begin{aligned}
C^{\text{Idenom}} &= 1 + C^{\text{Gpure}} + (k-1)\, C^{\text{Gcross}} && (61)\\
&= 1 + \left(\mu^2 \left(\frac{n}{n-1}\right)^2 + 2\mu \left(\frac{n}{n-1}\right)\right) \frac{(n-1)-(k-1)}{n} + 2\left(\mu^2 \epsilon^2 + 2\mu\epsilon\right) \frac{1}{n} && (62)
\end{aligned}
$$

Since $k \in \{1, \ldots, n\}$, thus, $k \le n$. With this and $n \ge 2$, obviously we obtain (62) > 0.

Thus, (35) is equivalent to (64) as follows:

$$
\begin{aligned}
C_k^{\text{cross}} C_k^{\text{numera}} - C_k^{\text{nutero}} \left(1 + C_k^{\text{pure}}\right) &> 0 && (63)\\
\Leftrightarrow \left(C_k^{\text{Idenom}}\right)^2 \left(C_k^{\text{cross}} C_k^{\text{numera}} - C_k^{\text{nutero}} \left(1 + C_k^{\text{pure}}\right)\right) &> 0 && (64)
\end{aligned}
$$

Thus, we will prove (64), an equivalent proposition of Lemma 2. By substituting (53), (54), (59), and (60) to (64), we obtain the following:

$$
\begin{aligned}
&\left(C_k^{\text{Idenom}}\right)^2 \left(C_k^{\text{cross}} C_k^{\text{numera}} - C_k^{\text{nutero}} \left(1 + C_k^{\text{pure}}\right)\right) && (65)\\
=\; &\left(C^{\text{Gcross}} - C^{\text{Gpure}} - 1\right)^2 \left(C^{\text{Gcross}} C^{\text{Gnumera}} - C^{\text{Gnutero}} C^{\text{Gpure}} - C^{\text{Gnutero}}\right). && (66)
\end{aligned}
$$

The transformation between (65) and (66) is checked by **SymPy**, a Computer Algebra System (CAS) designed for Python[1]. We also numerically checked the transformation by substituting random values and confirmed that the value matched. In (66), the proposition that (35) > 0 is equivalent to the second factor of (66), $C^{\text{Gcross}}C^{\text{Gnumera}} - C^{\text{Gnutero}}C^{\text{Gpure}} - C^{\text{Gnutero}} > 0$ since the first factor of (66) is obviously positive because it is squared. Thus, Lemma 2 is equivalent to

$$C^{\text{Gcross}}C^{\text{Gnumera}} - C^{\text{Gnutero}}C^{\text{Gpure}} - C^{\text{Gnutero}} > 0. \tag{67}$$

Thus, we will prove (67). By substituting (37), (36), (38), and (39) to $C^{\text{Gcross}}C^{\text{Gnumera}} - C^{\text{Gnutero}}C^{\text{Gpure}} - C^{\text{Gnutero}}$, we obtain:

$$\begin{aligned} & C^{\text{Gcross}}C^{\text{Gnumera}} - C^{\text{Gnutero}}C^{\text{Gpure}} - C^{\text{Gnutero}} & (68) \\ = \; & -\frac{1}{n(n-1)^7}\mu^2(\epsilon n - \epsilon - n)(\mu n + n - 1)^4 \\ & \Big(\epsilon^2\mu^4 n^2 + 2\epsilon^2\mu^3 n(n-1) + \epsilon^2\mu^2(n-1)^2 + 2\epsilon\mu^3 n^2 + 5\epsilon\mu^2 n(n-1) \\ & +4\epsilon\mu(n-1)^2 + \mu^2 n^2 + 4\mu n(n-1) + 4(n-1)^2\Big) & (69) \end{aligned}$$

We carefully examine (69). First, we focus on the denominator of (69). Since n is the number of a complete graph, by definition, $n \in \mathbb{N}$. However, the denominator $n(n-1)^7$ becomes 0 when $n = 1$. Thus, we have $n \in \{2, 3, 4, \ldots\}$.

Next, we focus on the numerator of (69). The first and third factor, μ^2 and $(\mu n + n - 1)^4$, respectively, is obviously positive. The last factor is also positive provided that $n \geq 2$ since $\epsilon > 0$ and $\mu > 0$.

$$\begin{aligned} & \epsilon^2\mu^4 n^2 + 2\epsilon^2\mu^3 n(n-1) + \epsilon^2\mu^2(n-1)^2 + 2\epsilon\mu^3 n^2 + 5\epsilon\mu^2 n(n-1) \\ & \quad +4\epsilon\mu(n-1)^2 + \mu^2 n^2 + 4\mu n(n-1) + 4(n-1)^2 > 0. \end{aligned} \tag{70}$$

Finally, the second factor in (69), $\epsilon n - \epsilon - n$, is negative under the condition that $\epsilon < \frac{n}{n-1}$. Remember that ϵ is a substitute of the 0 eigenvalue, and [1] and we set $\epsilon = 10^{-6}$. Thus, this condition is easily achieved because we usually set $0 < \epsilon << 1 < \frac{n}{n-1}$.

In summary, the conditions for (69) are:

$$\begin{aligned} n &\in \{2, 3, 4, \ldots\} & (71) \\ 0 &< \epsilon < 1 & (72) \end{aligned}$$

Under the conditions of (71) and (72), only the second factor in the numerator of (69) is negative, and the latter factors is positive. Thus, overall, (69) is positive under the conditions. Thus, Lemma 2 holds true.

□

References

[1] Quanquan Gu and Jiawei Han. Towards active learning on graphs: An error bound minimization approach. In *Proceedings of the IEEE International Conference on Data Mining (ICDM) 2012*, 2012.

[1] See `http://yoehara.com/` for details.