7. Proof of Lemma 2.2 – Local Packing Set

Towards the proof of Lemma 2.2, we develop a modified version of the Varshamov-Gilbert Lemma adapted to our specific model: the set of characteristic vectors of the S-T paths of a (p, k, d)-layer graph G.

Let $\delta_H(\mathbf{x}, \mathbf{y})$ denote the Hamming distance between two points $\mathbf{x}, \mathbf{y} \in \{0, 1\}^p$:

$$\delta_H(\mathbf{x}, \mathbf{y}) \triangleq |\{i : x_i \neq y_i\}|.$$

Lemma 7.4. *Consider a (p, k, d)-layer graph G on p vertices and the collection $\mathcal{P}(G)$ of S-T paths in G. Let*

$$\Omega \triangleq \big\{\mathbf{x} \in \{0, 1\}^p : \textit{supp}(\mathbf{x}) \in \mathcal{P}(G)\big\},$$

i.e., the set of characteristic vectors of all S-T paths in G. For every $\xi \in (0, 1)$, there exists a set, $\Omega_\xi \subset \Omega$ such that

$$\delta_H(\mathbf{x}, \mathbf{y}) > 2(1 - \xi) \cdot k, \quad \forall \mathbf{x}, \mathbf{y} \in \Omega_\xi, \mathbf{x} \neq \mathbf{y}, \tag{29}$$

and

$$\log |\Omega_\xi| \geq \log \tfrac{p-2}{k} + (\xi \cdot k - 1) \cdot \log d - k \cdot H(\xi), \tag{30}$$

where $H(\cdot)$ is the binary entropy function.

Proof. Consider a labeling $1, \ldots, p$ of the p vertices in G, such that variable ω_i is associated with vertex i. Each point $\boldsymbol{\omega} \in \Omega$ is the characteristic vector of a set in $\mathcal{P}(G)$; nonzero entries of $\boldsymbol{\omega}$ correspond to vertices along an S-T path in G. With a slight abuse of notation, we refer to $\boldsymbol{\omega}$ as a path in G. Due to the structure of the (p, k, d)-layer graph G, all points in Ω have exactly $k + 2$ nonzero entries, *i.e.*,

$$\delta_H(\boldsymbol{\omega}, \mathbf{0}) = k + 2, \qquad \forall \boldsymbol{\omega} \in \Omega.$$

Each vertex in $\boldsymbol{\omega}$ lies in a distinct layer of G. In turn, for any pair of points $\boldsymbol{\omega}, \boldsymbol{\omega}' \in \Omega$,

$$\delta_H(\boldsymbol{\omega}, \boldsymbol{\omega}') = 2 \cdot \big(k - |\{i : \omega_i = \omega_i' = 1\}| - 2\big). \tag{31}$$

Note that the Hamming distance between the two points is a linear function of the number of their common nonzero entries, while it can take only even values with a maximum value of $2k$.

Without loss of generality, let S and T corresponding to vertices 1 and p, respectively. Then, the above imply that

$$\omega_1 = \omega_p = 1, \quad \forall \boldsymbol{\omega} \in \Omega.$$

Consider a fixed point $\widehat{\boldsymbol{\omega}} \in \Omega$, and let $\mathcal{B}(\widehat{\boldsymbol{\omega}}, r)$ denote the Hamming ball of radius r centered at $\widehat{\boldsymbol{\omega}}$, *i.e.*,

$$\mathcal{B}(\widehat{\boldsymbol{\omega}}, r) \;\triangleq\; \big\{\boldsymbol{\omega} \in \{0, 1\}^p : \delta_H(\widehat{\boldsymbol{\omega}}, \boldsymbol{\omega}) \leq r\big\}.$$

The intersection $\mathcal{B}(\widehat{\boldsymbol{\omega}}, r) \cap \Omega$ corresponds to S-T paths in G that have at least $k - r/2$ additional vertices in common with $\widehat{\boldsymbol{\omega}}$ besides vertices 1 and p that are common to all paths in Ω:

$$\begin{aligned}
&\mathcal{B}(\widehat{\boldsymbol{\omega}}, r) \cap \Omega \\
&= \{\boldsymbol{\omega} \in \Omega : \delta_H(\widehat{\boldsymbol{\omega}}, \boldsymbol{\omega}) \leq r\} \\
&= \big\{\boldsymbol{\omega} \in \Omega : |\{i : \widehat{\omega}_i = \omega_i = 1\}| \geq k - \tfrac{r}{2} + 2\big\},
\end{aligned}$$

where the last equality is due to (31). In fact, due to the structure of G, the set $\mathcal{B}(\widehat{\boldsymbol{\omega}}, r) \cap \Omega$ corresponds to the S-T paths that *meet* $\widehat{\boldsymbol{\omega}}$ in at least $k - r/2$ intermediate layers. Taking into account that $|\Gamma_{\text{in}}(v)| = |\Gamma_{\text{out}}(v)| = d$, for all vertices v in $V(G)$ (except those in the first and last layer),

$$|\mathcal{B}(\widehat{\boldsymbol{\omega}}, r) \cap \Omega| \leq \binom{k}{k - \frac{r}{2}} \cdot d^{k - \left(k - \frac{r}{2}\right)} = \binom{k}{k - \frac{r}{2}} \cdot d^{\frac{r}{2}}.$$

Now, consider a *maximal* set $\Omega_\xi \subset \Omega$ satisfying (29), *i.e.*, a set that cannot be augmented by any other point in Ω. The union of balls $\mathcal{B}(\boldsymbol{\omega}, 2(1 - \xi) \cdot (k - 1))$ over all $\boldsymbol{\omega} \in \Omega_\xi$ covers Ω. To verify that, note that if there exists $\boldsymbol{\omega}' \in \Omega \backslash \Omega_\xi$ such that $\delta_H(\boldsymbol{\omega}, \boldsymbol{\omega}') > 2(1 - \xi) \cdot (k - 1)$, $\forall \boldsymbol{\omega} \in \Omega_\xi$, then $\Omega_\xi \cup \{\boldsymbol{\omega}'\}$ satisfies (29) contradicting the maximality of Ω_ξ. Based on the above,

$$\begin{aligned}
|\Omega| &\leq \sum_{\boldsymbol{\omega} \in \Omega_\xi} |\mathcal{B}(\boldsymbol{\omega}, 2(1 - \xi) \cdot k) \cap \Omega| \\
&\leq \sum_{\mathbf{x} \in \Omega_\xi} \tbinom{k}{k - (1 - \xi)k} \cdot d^{(1 - \xi) \cdot k} \\
&\leq \sum_{\mathbf{x} \in \Omega_\xi} \tbinom{k}{\xi k} \cdot d^{(1 - \xi) \cdot k} \\
&\leq |\Omega_\xi| \cdot 2^{k \cdot H(\xi)} \cdot d^{(1 - \xi) \cdot k}.
\end{aligned}$$

Taking into account that

$$|\Omega| = |\mathcal{P}(G)| = \frac{p - 2}{k} \cdot d^{k - 1},$$

we conclude that

$$\frac{p - 2}{k} \cdot d^{k - 1} \leq |\Omega_\xi| \cdot 2^{k \cdot H(\xi)} \cdot d^{(1 - \xi) \cdot k},$$

from which the desired result follows. □

Lemma 2.2. (Local Packing) *Consider a (p, k, d)-layer graph G on p vertices with $k \geq 4$ and $\log d \geq 4 \cdot H(3/4)$. For any $\epsilon \in (0, 1]$, there exists a set $\mathcal{X}_\epsilon \subset \mathcal{X}(G)$ such that*

$$\epsilon / \sqrt{2} \;<\; \|\mathbf{x}_i - \mathbf{x}_j\|_2 \;\leq\; \sqrt{2} \cdot \epsilon,$$

for all $\mathbf{x}_i, \mathbf{x}_j \in \mathcal{X}_\epsilon$, $\mathbf{x}_i \neq \mathbf{x}_j$, and

$$\log |\mathcal{X}_\epsilon| \;\geq\; \log \frac{p - 2}{k} + \frac{1}{4} \cdot k \log d.$$

Proof. Without loss of generality, consider a labeling $1, \ldots, p$ of the p vertices in G, such that S and T correspond to vertices 1 and p, respectively. Let

$$\Omega \triangleq \left\{ \mathbf{x} \in \{0,1\}^p : \ \mathrm{supp}(\mathbf{x}) \in \mathcal{P}(G) \right\},$$

where $\mathcal{P}(G)$ is the collection of S-T paths in G. By Lemma 7.4, and for $\xi = 3/4$, there exists a set $\Omega_\xi \subseteq \Omega$ such that

$$\delta_H(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j) > \frac{1}{2} \cdot k, \tag{32}$$

$\forall \boldsymbol{\omega}_i, \boldsymbol{\omega}_j \in \Omega_\xi$, $\boldsymbol{\omega}_i \neq \boldsymbol{\omega}_j$, and,

$$\begin{aligned} \log|\Omega_\xi| &\geq \log \tfrac{p-2}{k} + \left(\tfrac{3}{4} \cdot k - 1\right) \log d - k \cdot H\left(\tfrac{3}{4}\right) \\ &\geq \log \tfrac{p-2}{k} + \tfrac{2}{4} \cdot k \cdot \log d - k \cdot H\left(\tfrac{3}{4}\right) \\ &\geq \log \tfrac{p-2}{k} + \tfrac{1}{4} \cdot k \cdot \log d \end{aligned} \tag{33}$$

where the second and third inequalites hold under the assumptions of the lemma; $k \geq 4$ and $\log d \geq 4 \cdot H(3/4)$.

Consider the bijective mapping $\psi : \Omega_\xi \to \mathbb{R}^p$ defined as

$$\psi(\boldsymbol{\omega}) = \left[\sqrt{\tfrac{(1-\epsilon^2)}{2}} \cdot \omega_1, \ \tfrac{\epsilon}{\sqrt{k}} \cdot \boldsymbol{\omega}_{2:p-1}, \ \sqrt{\tfrac{(1-\epsilon^2)}{2}} \cdot \omega_p \right].$$

We show that the set

$$\mathcal{X}_\epsilon \triangleq \left\{ \psi(\boldsymbol{\omega}) : \ \boldsymbol{\omega} \in \Omega_\xi \right\}.$$

has the desired properties. First, to verify that $\mathcal{X}_\epsilon$ is a subset of $\mathcal{X}(G)$, note that $\forall \boldsymbol{\omega} \in \Omega_\xi \subset \Omega$,

$$\mathrm{supp}(\psi(\boldsymbol{\omega})) = \mathrm{supp}(\boldsymbol{\omega}) \in \mathcal{P}(G), \tag{34}$$

and

$$\|\psi(\boldsymbol{\omega})\|_2^2 = 2 \cdot \frac{(1-\epsilon^2)}{2} + \frac{\epsilon^2}{k} \cdot \sum_{i=2}^{p-1} \omega_i = 1.$$

Second, for all pairs of points $\mathbf{x}_i, \mathbf{x}_j \in \mathcal{X}_\epsilon$,

$$\|\mathbf{x}_i - \mathbf{x}_j\|_2^2 = \delta_H(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j) \cdot \frac{\epsilon^2}{k} \leq 2 \cdot k \cdot \frac{\epsilon^2}{k} = 2 \cdot \epsilon^2.$$

The inequality follows from the fact that $\delta_H(\boldsymbol{\omega}, \mathbf{0}) = k+2$ $\omega_1 = 1$ and $\omega_p = 1$, $\forall \boldsymbol{\omega} \in \Omega_\xi$, and in turn

$$\delta_H(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j) \leq 2 \cdot k.$$

Similarly, for all pairs $\mathbf{x}_i, \mathbf{x}_j \in \mathcal{X}_\epsilon$, $\mathbf{x}_i \neq \mathbf{x}_j$,

$$\|\mathbf{x}_i - \mathbf{x}_j\|_2 = \delta_H(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j) \cdot \frac{\epsilon^2}{k} \geq \frac{1}{2} \cdot k \cdot \frac{\epsilon^2}{k} = \frac{\epsilon^2}{2},$$

where the inequality is due to (32). Finally, the lower bound on the cardinality of $\mathcal{X}_\epsilon$ follows immediately from (33) and the fact that $|\mathcal{X}_\epsilon| = |\Omega_\xi|$, which completes the proof. □

8. Details in proof of Lemma 1

We want to show that if

$$\epsilon^2 = \min\left\{ 1, \ \frac{C' \cdot (1+\beta)}{\beta^2} \cdot \frac{\log \frac{p-2}{k} + \frac{k}{4} \cdot \log d}{n} \right\},$$

for an appropriate choice of $C' > 0$, then the following two conditions (Eq. (13)) are satisfied:

$$n \cdot \frac{2\epsilon^2 \beta^2}{(1+\beta)} \frac{1}{\log|\mathcal{X}_\epsilon|} \leq \frac{1}{4} \ \text{ and } \ \log|\mathcal{X}_\epsilon| \geq 4 \log 2.$$

For the second inequality, recall that by Lemma 2.2,

$$\log|\mathcal{X}_\epsilon| \ \geq \ \log \frac{p-2}{k} + \frac{1}{4} \cdot k \log d \ > \ 0. \tag{35}$$

Under the assumptions of Thm. 1 on the parameters k and d (note that $p - 2 \geq k \cdot d$ by the structure of G),

$$\log|\mathcal{X}_\epsilon| \geq \log \frac{p-2}{k} + \frac{k}{4} \cdot \log d \geq 4 \cdot H(3/4) \geq 4 \log 2,$$

which is the desired result.

For the first inequality, we consider two cases:

- First, we consider the case where $\epsilon^2 = 1$, *i.e.*,

$$\epsilon^2 = 1 \leq \frac{C' \cdot (1+\beta)}{\beta^2} \cdot \frac{\log \frac{p-2}{k} + \frac{k}{4} \cdot \log d}{n}.$$

  Equivalently,

$$n \cdot \frac{2\epsilon^2 \beta^2}{(1+\beta)} \leq 2 \cdot C' \cdot \left(\log \frac{p-2}{k} + \frac{k}{4} \cdot \log d \right). \tag{36}$$

- In the second case,

$$\epsilon^2 = \frac{C' \cdot (1+\beta)}{\beta^2} \cdot \frac{\log \frac{p-2}{k} + \frac{k}{4} \cdot \log d}{n},$$

  which implies that

$$n \cdot \frac{2\epsilon^2 \beta^2}{(1+\beta)} = 2 \cdot C' \cdot \left(\log \frac{p-2}{k} + \frac{k}{4} \cdot \log d \right). \tag{37}$$

Combining (36) or (37), with (35), we obtain

$$n \cdot \frac{2\epsilon^2 \beta^2}{(1+\beta)} \frac{1}{\log|\mathcal{X}_\epsilon|} \ \leq \ 2 \cdot C' \ \leq \ \frac{1}{4}$$

for $C' \leq 1/8$.

We conclude that for ϵ chosen as in (12), the conditions in (13) hold.

9. Other

Assumption 1. *There exist i.i.d. random vectors* $\mathbf{z}_1, \dots, \mathbf{z}_n \in \mathbb{R}^p$*, such that* $\mathbb{E}\mathbf{z}_i = \mathbf{0}$ *and* $\mathbb{E}\mathbf{z}_i\mathbf{z}_i^\top = \mathbb{I}_p$*,*

$$\mathbf{y} = \mu + \mathbf{\Sigma}^{1/2}\mathbf{z}_i \tag{38}$$

and

$$\sup_{\mathbf{x} \in \mathbb{S}_2^{p-1}} \|\mathbf{z}_i^\top \mathbf{x}\|_{\psi_2} \leq K, \tag{39}$$

where $\mu \in \mathbb{R}^p$ *and* $K > 0$ *is a constant depending on the distribution of* $\mathbf{z}_i$*s.*