Theorem 3: *Leveraging multiple subgraphs enhances entropy and improves label discrimination.* Specifically,

$$P_{\text{multi}}^{(t)} \geq P_{\text{single}}^{(t)}, \quad H_{\text{multi}}^{(t)} \geq H_{\text{single}}^{(t)}, \tag{1}$$

where $P_{\text{single}}^{(t)}$ and $H_{\text{single}}^{(t)}$ represent the probability and entropy of label conflicts for a single subgraph, while $P_{\text{multi}}^{(t)}$ and $H_{\text{multi}}^{(t)}$ represent the corresponding quantities for multiple subgraphs.

Proof

The probability of label conflict for a single subgraph at the t-th update is:

$$P_{\text{single}}^{(t)} = \frac{\text{Conflicting pairs}_{\text{single}}^{(t)}}{\binom{n}{2}}, \tag{2}$$

where Conflicting pairs$_{\text{single}}^{(t)}$ is the number of conflicting pairs in a single subgraph.

Let there be k subgraphs. The aggregated label conflict probability across multiple subgraphs is:

$$P_{\text{multi}}^{(t)} = \frac{\text{Conflicting pairs}_{\text{multi}}^{(t)}}{\binom{n}{2}}, \tag{3}$$

where Conflicting pairs$_{\text{multi}}^{(t)}$ accounts for the combined label diversity across all subgraphs.

It is easy to find that multiple subgraphs increase label diversity through diverse sampling strategies, reducing the chance of label conflicts:

$$P_{\text{multi}}^{(t)} \geq P_{\text{single}}^{(t)}. \tag{4}$$

The entropy for a single subgraph is given by:

$$H_{\text{single}}^{(t)} = -\sum_{x \in \mathcal{H}} P_{\text{single}}^{(t)}(x) \log P_{\text{single}}^{(t)}(x), \tag{5}$$

where $P_{\text{single}}^{(t)}(x)$ is the probability distribution of labels in the single subgraph.

The entropy for multiple subgraphs combines the diversity across k subgraphs are defined as follows:

$$H_{\text{multi}}^{(t)} = -\sum_{x \in \mathcal{H}} P_{\text{multi}}^{(t)}(x) \log P_{\text{multi}}^{(t)}(x), \tag{6}$$

where $P_{\text{multi}}^{(t)}(x)$ represents the aggregated label distribution across all subgraphs.

As multiple subgraphs reduce label conflicts and increase diversity, the entropy satisfies:

$$H_{\text{multi}}^{(t)} > H_{\text{single}}^{(t)}. \tag{7}$$

From (4) and (7), we establish:

$$P_{\text{multi}}^{(t)} \geq P_{\text{single}}^{(t)}, \quad H_{\text{multi}}^{(t)} > H_{\text{single}}^{(t)}. \tag{8}$$