A. Proof of Theorem 1

Proof 1 *By virtue of the relationship between mutual information and information entropy, we obtain:*

$$I(S(\boldsymbol{x})^{(l)}; H(\boldsymbol{x})^{(l)}) = \\ \mathrm{H}(S(\boldsymbol{x})^{(l)}) + \mathrm{H}(H(\boldsymbol{x})^{(l)}) - \mathrm{H}(S(\boldsymbol{x})^{(l)}, H(\boldsymbol{x})^{(l)}), \tag{10}$$

where $\mathrm{H}(\cdot)$ *is the information entropy and* $\mathrm{H}(\cdot,\cdot)$ *is the joint entropy. And the information gain or Kullback-Leibler divergence with information entropy is defined as:*

$$D_{KL}(S(\boldsymbol{x})^{(l)} \| H(\boldsymbol{x})^{(l)}) = \\ \mathrm{H}(S(\boldsymbol{x})^{(l)}, H(\boldsymbol{x})^{(l)}) - \mathrm{H}(S(\boldsymbol{x})^{(l)}). \tag{11}$$

Combining Eq. (10) *with* (11)*, we get:*

$$\begin{aligned} &I(S(\boldsymbol{x})^{(l)}; H(\boldsymbol{x})^{(l)}) = \mathrm{H}(S(\boldsymbol{x})^{(l)}) + \mathrm{H}(H(\boldsymbol{x})^{(l)}) \\ &- D_{KL}(S(\boldsymbol{x})^{(l)} \| H(\boldsymbol{x})^{(l)}) - \mathrm{H}(S(\boldsymbol{x})^{(l)}) \\ &= \mathrm{H}(H(\boldsymbol{x})^{(l)}) - D_{KL}(S(\boldsymbol{x})^{(l)} \| H(\boldsymbol{x})^{(l)}). \end{aligned} \tag{12}$$

In terms of Eq. (12)*, we observe that* $I(S(\boldsymbol{x})^{(l)}; H(\boldsymbol{x})^{(l)})$ *is negatively correlated with* $D_{KL}(S(\boldsymbol{x})^{(l)} \| H(\boldsymbol{x})^{(l)})$*. Associated with* $D_{KL}(S(\boldsymbol{x})^{(l)} \| H(\boldsymbol{x})^{(l)}) \sim \delta_f^{(l)}$ [18]*, maximizing* $I(S(\boldsymbol{x})^{(l)}; H(\boldsymbol{x})^{(l)})$ *essentially minimizes* $D_{KL}(S(\boldsymbol{x})^{(l)} \| H(\boldsymbol{x})^{(l)})$ *and* $\delta_f^{(l)}$*, which attains the goal of graph smoothing:*

$$\begin{aligned} I(S(\boldsymbol{x})^{(l)}; H(\boldsymbol{x})^{(l)}) &\sim \frac{1}{D_{KL}(S(\boldsymbol{x})^{(l)} \| H(\boldsymbol{x})^{(l)})} \\ &\sim \frac{1}{\delta_f^{(l)}}. \end{aligned} \tag{13}$$

B. Subgraph Statistics

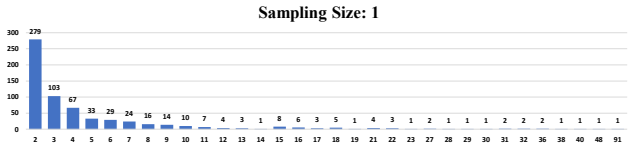

Figure 6. Subgraph statistics on Citeseer by using TAPS strategy to perform the subgraph partition with sampling size 1.

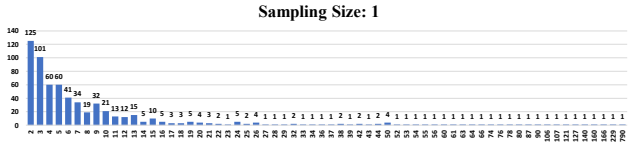

Figure 7. Subgraph statistics on Amazon Photo by using TAPS strategy to perform the subgraph partition with sampling size 1.

Subgraph statistics on Citeseer (Fig. 6), Amazon Photo (Fig. 7), Coauthor CS (Fig. 8), Coauthor Physics (Fig. 9), and Pubmed (Fig. 10) by using TAPS strategy to perform the subgraph partition with sampling size 1.

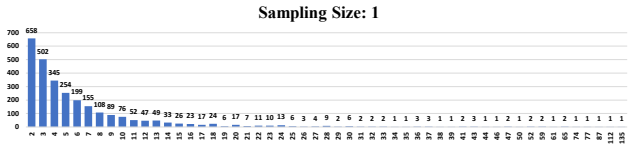

Figure 8. Subgraph statistics on Coauthor CS by using TAPS strategy to perform the subgraph partition with sampling size 1.

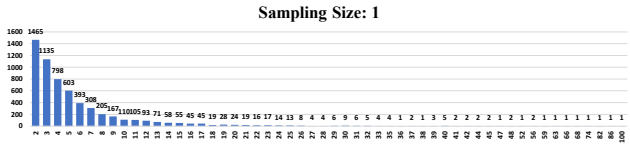

Figure 9. Subgraph statistics on Coauthor Physics by using TAPS strategy to perform the subgraph partition with sampling size 1.

C. Subgraph Partition Visualization

Visualization of part of the subgraphs derived from TAPS with sampling size 1 on Pubmed (Fig. 11), Citeseer (Fig. 12), Amazon Photo (Fig. 13), Coauthor CS (Fig. 14), and Coauthor Physics (Fig. 15).

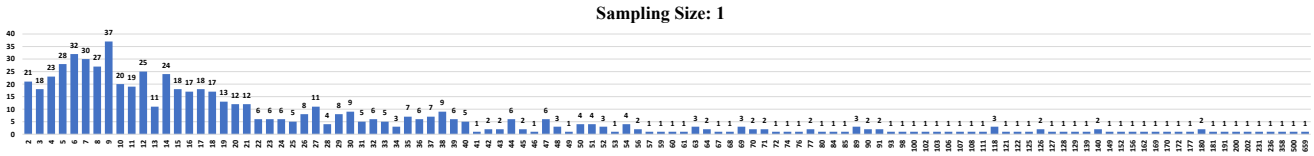

Figure 10. Subgraph statistics on Pubmed by using TAPS strategy to perform the subgraph partition with sampling size 1.

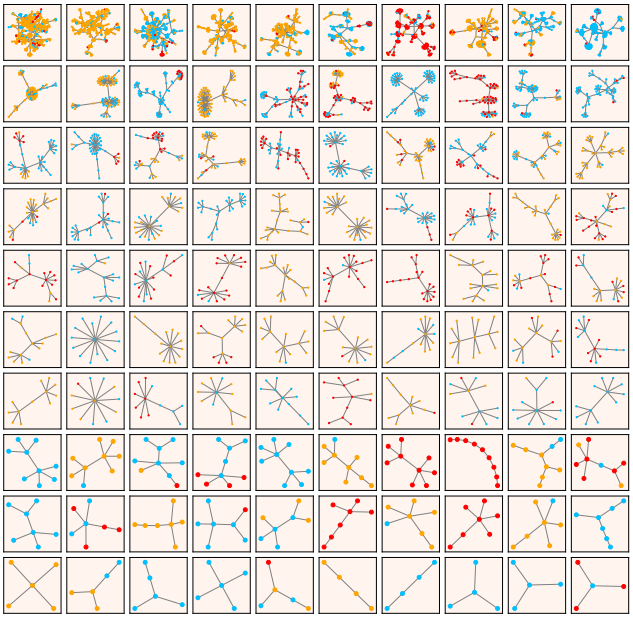

Figure 11. Visualization of part of the subgraphs derived from TAPS with sampling size 1 on Pubmed.

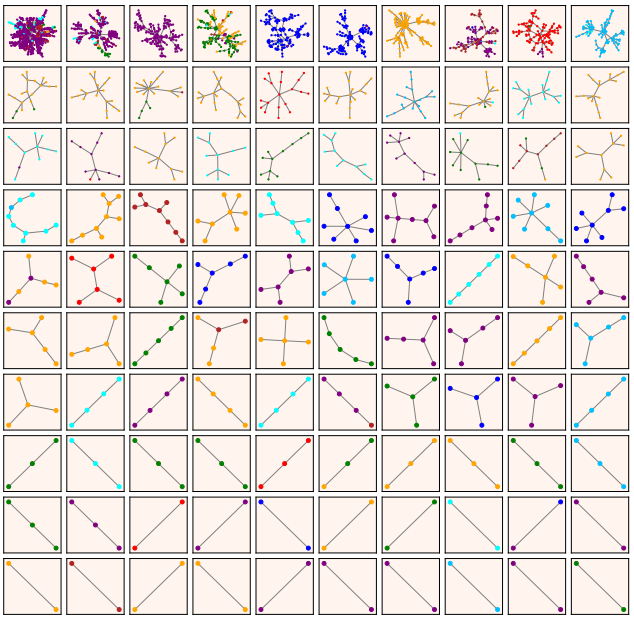

Figure 13. Visualization of part of the subgraphs derived from TAPS with sampling size 1 on Amazon Photo.

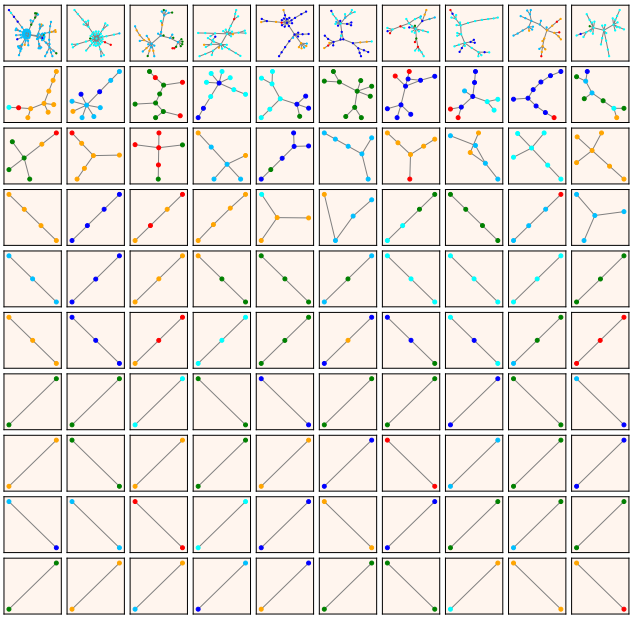

Figure 12. Visualization of part of the subgraphs derived from TAPS with sampling size 1 on Citeseer.

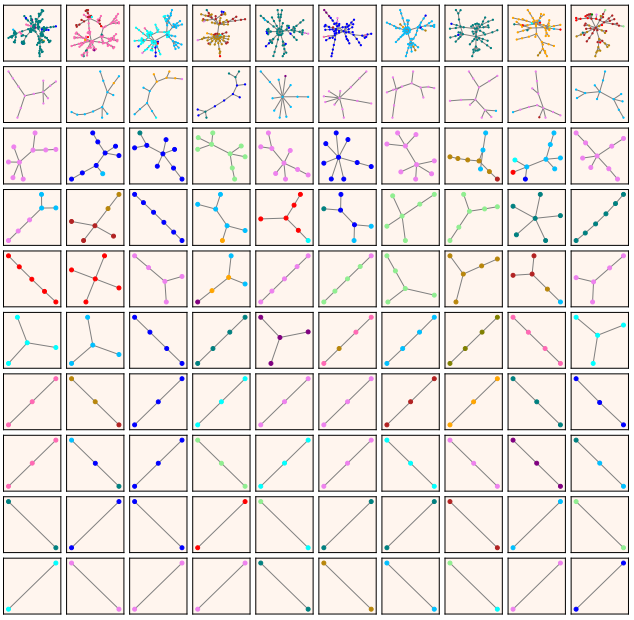

Figure 14. Visualization of part of the subgraphs derived from TAPS with sampling size 1 on Coauthor CS.

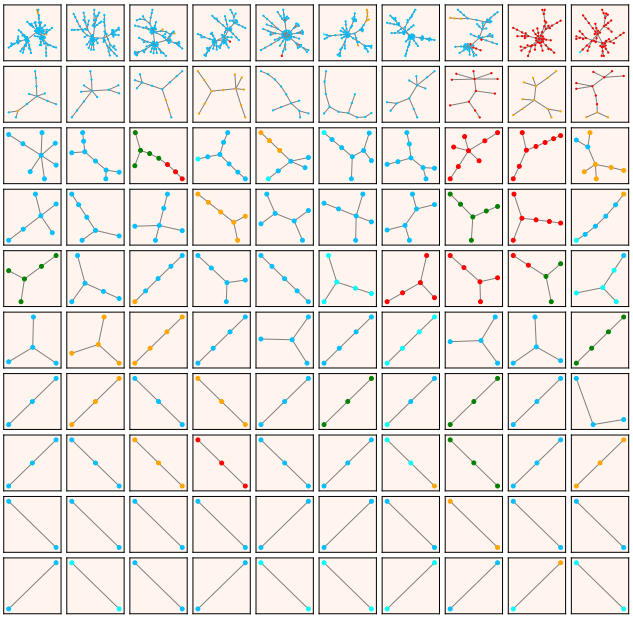

Figure 15. Visualization of part of the subgraphs derived from TAPS with sampling size 1 on Coauthor Physics.