## APPENDIX

Proposition 1. Let $v$ be an arbitrary target node. In the vanilla scheme, the embedding of $\mathbf{h}_v^{(\text{vanilla})}$ is computed as,

$$\mathbf{h}_v^{(l)} = \text{AGG}_{\text{all}}^{(l)}\left(\left\{\text{AGG}_r^{(l)}\left(\left\{\mathbf{h}_u^{(l-1)}, u \in N_r(v)\right\}\right), r \in \mathcal{R}\right\}\right)$$

Under RAF, the embedding of $\mathbf{h}_v^{(\text{RAF})}$ is computed through two steps. First, by lines 2-6 in Algorithm 1, the intermediate embedding of the node from different relations is computed remotely and aggregated at the remote worker. This yields a set of partial aggregations,

$$\left\{\text{AGG}_r^{(l)}\left(\left\{\mathbf{h}_u^{(l-1)}, u \in N_r(v)\right\}\right), r \in \mathcal{R}\right\}$$

Then, these messages are sent to the worker with the target node, and the worker performs cross-relation aggregation $\text{AGG}_{\text{all}}^{(l)}$ on these partial aggregations (lines 8-12 in Algorithm 1). Since AGG(.) is order (permutation) invariant, we have that

$$\mathbf{h}_v^{(\text{RAF})} = \mathbf{h}_v^{(\text{vanilla})}$$

□

Proposition 2. Given an arbitrary partition $G_1, G_2$, we first focus our analysis on $G_1$. Let $k$ denote the number of relations in the graphs, and recall that $\text{B}(G_1)$ is the number of boundary node in $G_1$. Let $v$ be an arbitrary node in $G_1$ that requires communication from $G_2$. By definition, $v$ must be a boundary node as it has a neighborhood from partition $G_2$. By lines 4-12 in Algorithm 1, only the intermediate representation of each relation $r$ in $v$'s neighborhood needs to be communicated. Therefore, the number of messages received for $v$ in $G_1$ is at most $k$. Therefore, the number of messages received for $G_1$ from $G_2$ is at most

$$k\text{B}(G_1).$$

Since $k$ is a constant for a given dataset or task and is independent of graph size, we have the communication complexity of worker $G_1$ is

$$\Theta(\text{B}(G_1)).$$

This completes the proof since the analysis for $G_2$ is symmetric.

□

Proposition 3. Again, let us focus on the analysis in $G_1$. By definition of boundary node, a boundary node in $G_1$ must be an endpoint of at least one cross-partition edge. However, there may be multiple cross-partition edges connecting to the node. Therefore, we have that

$$\text{B}(G_1) \leq \text{E}(G_1, G_2).$$

By a symmetric argument, we have that

$$\text{B}(G_2) \leq \text{E}(G_1, G_2).$$

Combining the result above, we can obtain that

$$\max\{\text{B}(G_1), \text{B}(G_2)\} \leq \text{E}(G_1, G_2).$$

□