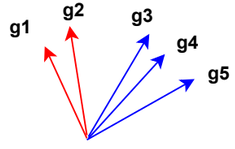
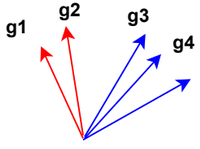
Local model updates g1,g2,g3,g4,g5

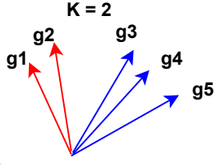


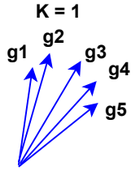
Compute mutual cosine distances



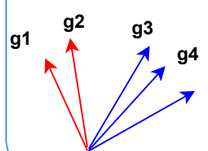
$$c_{ij}^{t} = 1 - \left(\frac{\langle g_i, g_j \rangle}{\|g_i\|_2 \cdot \|g_j\|_2} \right)$$

Cluster the cosine distance with K-means which K is determined by Gap-statistics





Aggregation: When K is 2, weight the aggregation towards the larger cluster. Otherwise, perform weighted aggregation of all model updates.



$$g = \frac{1}{(\beta^{l_3} + \beta^{l_4} + \beta^{l_5})} \left(\beta^{l_3} * g_3 + \beta^{l_4} * g_4 + \beta^{l_5} * g_5 \right)$$