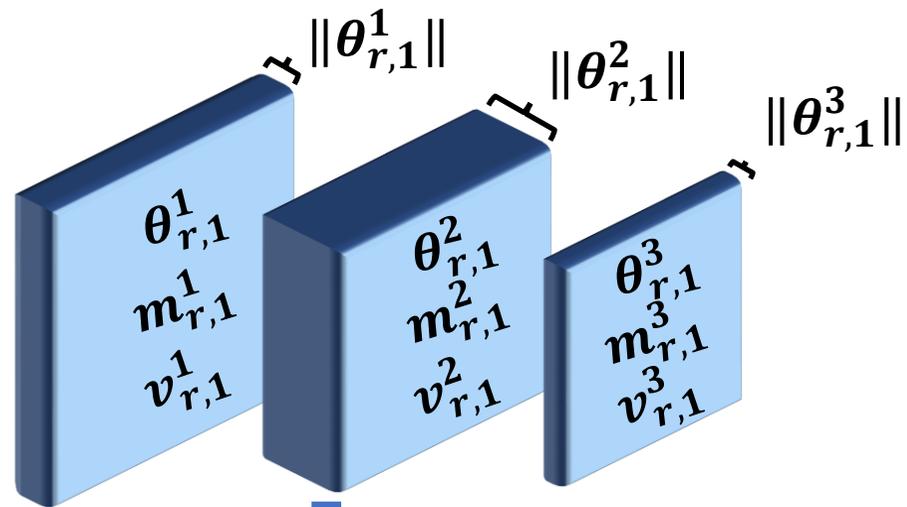
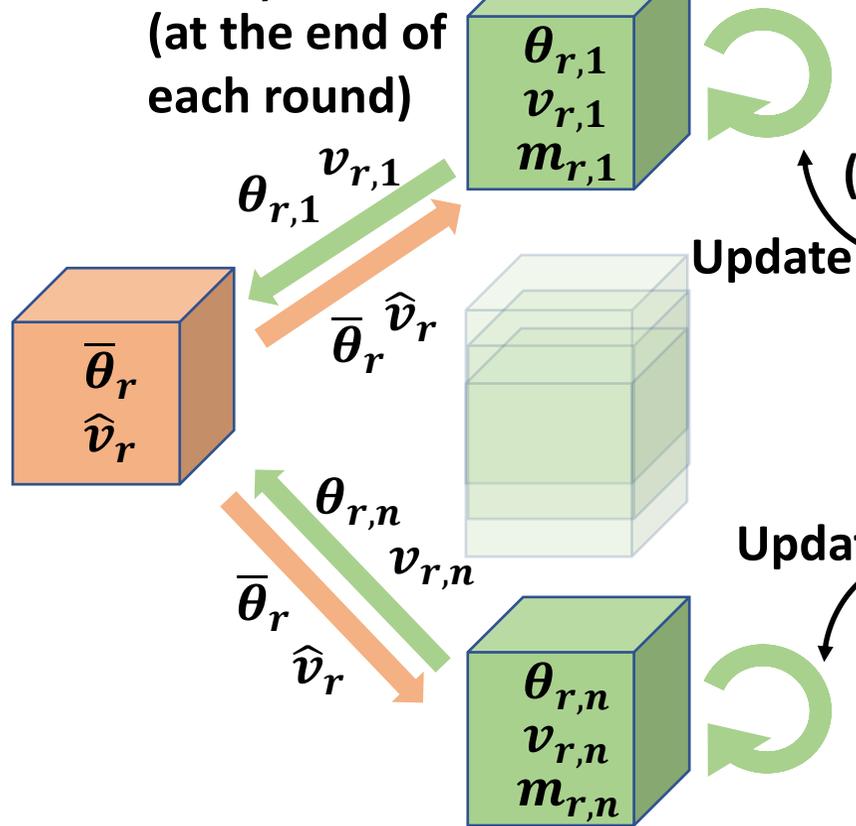


Global Aggregation

Local updates



Adaptive ratio	$p_{r,i}^1$	$p_{r,i}^2$	$p_{r,i}^3$
Layerwise Normalization	$\frac{\ \theta_{r,i}^1\ }{\ p_{r,i}^1\ } p_{r,i}^1$	$\frac{\ \theta_{r,i}^2\ }{\ p_{r,i}^2\ } p_{r,i}^2$	$\frac{\ \theta_{r,i}^3\ }{\ p_{r,i}^3\ } p_{r,i}^3$

